IFW

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/929,356	08/15/2001	Tameo Yanagino	107101-00036	8870

7590 06/30/2005

NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
Metropolitan Square
Suite 330 - G Street Lobby
655 Fifteenth Street, N.W.
Washington, DC 20005-5701

EXAMINER	
GRAYSAY, TAMARA L	

ART UNIT	PAPER NUMBER
3623	

DATE MAILED: 06/30/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

RECEIVED
OIPE/IAP
AUG 0 8 2005

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)	
	09/929,356	YANAGINO ET AL.	
	Examiner	Art Unit	
	Tamara L. Graysay	3623	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-22 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-22 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on 15 August 2001 is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☒ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date (2 pages).

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

Priority

1. Receipt is acknowledged of papers submitted under 35 U.S.C. 119(a)-(d), which papers have been placed of record in the file.

Oath/Declaration

2. The oath or declaration is defective. A new oath or declaration in compliance with 37 CFR 1.67(a) identifying this application by application number and filing date is required. See MPEP §§ 602.01 and 602.02.

The oath or declaration is defective because:

Although it identifies the foreign application, it does not state that the foreign application had been filed by the inventor(s) or by the assignee, or the legal representative or agent, of the inventor, or on behalf of the inventor, pursuant to MPEP § 201.13, II, C.

Specification

3. The lengthy specification has not been checked to the extent necessary to determine the presence of all possible minor errors. Applicant's cooperation is requested in correcting any errors of which applicant may become aware in the specification.

4. The disclosure is objected to because of the following informalities:

a. Page 9, lines 3-8, appears to be inconsistent as to the explanations of the ratio. The first explanation of the number of orders ratio = the number of orders after orders

were nil over the number of orders before orders were nil. Whereas the second explanation of the ratio is the number of orders before the order expired and after the order expired. It seems that the second explanation is the inverse of the first explanation and not the same ratio. Clarification is required in response to this Office action.

b. Page 11, line 6, CRT should be spelled out at its first occurrence.

c. Page 13, line 20, "form" should be from.

Appropriate correction is required.

Claim Objections

5. A series of singular dependent claims is permissible in which a dependent claim refers to a preceding claim which, in turn, refers to another preceding claim.

A claim which depends from a dependent claim should not be separated by any claim which does not also depend from said dependent claim. It should be kept in mind that a dependent claim may refer to any preceding independent claim. In general, applicant's sequence will not be changed. See MPEP § 608.01(n).

In the present application, the examiner has not changed the numbering or sequence of the claims.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

6. Claims 1-22 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claims 1, 7, 12, and 18, line 5, "the predetermined level" lacks antecedent basis in the claim.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

7. Claims 1-22 are rejected under 35 U.S.C. 101 because the claimed invention is directed to nonstatutory subject matter.

Claims 1-11

A process claim is statutory if it is limited to a practical application within the technological arts.

First, a practical application is one that is useful, concrete, and tangible, i.e., a real world value and reproducible. Although the claims include forecasting a number of orders, an inherently useful result, a process that merely manipulates an abstract idea or performs a purely mathematical algorithm is nonstatutory despite the fact that it might inherently have some usefulness. In re Sarkar, 200 USPQ 132, 139 (CCPA 178). In the present application, the steps of gathering data and performing the steps of the equation(s) to arrive at the forecasted number of orders is nothing more than gathering and substituting values in an equation dictated by the

mathematical formula. Such steps are nonstatutory, without producing something that is concrete and tangible.

Second, a process to be performed upon subject matter to be transformed and reduce to a different state or thing is within the technological arts. A process encompassing statutory subject matter is one that requires that certain things be done with certain substances, and in a certain order; but the tools to be used in doing this may be of secondary consequence. MPEP § 2106,IV,B,2. In the present application, the claims are drawn to a method of compiling data and performing mathematical steps. Each step, as claimed, is a human activity not performed on any subject matter to be transformed and reduced to a different state, the claims are therefore nonstatutory, i.e., not within the technological arts.

Therefore, process claims 1-11 are directed to nonstatutory subject matter.

Claims 12-22

The claims are drawn to a system for forecasting comprising "means" for performing functions. Looking to the specification for a determination of the scope of the claimed "means" reveals no particular machine or manufacture for performing the claimed functions. The disclosure does not include a computer program or logic circuit in support of the "means" limitation in the claim. Product claims not drawn to a specific machine or manufacture are evaluated as to the process to be performed by the product. Thus, the claims are evaluated as to the process to be performed by the system or product. Claims to processes that do nothing more than solve mathematical problems or manipulate abstract ideas or concepts are nonstatutory. If the steps of the claimed process manipulate only numbers, abstract concepts or ideas, or signals representative thereof, then the process does not manipulate appropriate subject matter. In the

present application, system claims 12-22 are nonstatutory because only an inherent usefulness has been claimed (as noted with regard to process claims 1-11 above) without producing something that is concrete and tangible. Further, the system process, as claimed, is a human activity not performed on any subject matter to be transformed and reduced to a different state, the claims are therefore nonstatutory, i.e., not within the technological arts.

Therefore, system claims 12-22 are directed to nonstatutory subject matter.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 1-6, 10, 12-17 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Mujtaba (article, Enterprise modeling and simulation: complex dynamic behavior of a simple model of manufacturing).

Claims 1- 10

Mujtaba discloses time analysis of orders for products (e.g., Adder-1, Adder-2, Adder-3, Adde-4) and compares the results of enterprise modeling and simulation for each of the products (for example, page 101, figures 15, 16 which depict orders after expiration represented by the curved line for each Adder/product). Mujtaba also mentions Monte Carlo simulation for analysis of business-oriented economics (page 82, left column, lines 1-11 a company finance department was modeled and the model subject to Monte Carlo simulation analysis). Further, Mujtaba

teaches in figure 4 the inventory curve for a product life cycle. Mujtaba teaches (claims 2-4, 6) forecasting for products having different order rates (page 92, right column, third and fourth paragraphs) and the effects of time series (page 93, figure 8). Mujtaba also teaches (claim 10) checking accuracy of forecast number of orders and making changes based on the comparison (page 91-92, A/F ratio, comparing forecasted demand with actual demand as they relate to product orders). Mujtaba teaches (claim 10) using initial orders as early indicators of a life cycle and revising the forecast after a certain period of time (page 94, right column, second paragraph). Mujtaba suggests (claim 7) combining probability distribution and simulation (page 99, right column, fifth paragraph, simulation and least sum of squares analysis). Mujtaba infers, rather than explicitly disclose, that products are categorized (page 101, right column, second paragraph, part commonality).

The examiner takes Official notice that it is within the level of ordinary skill in the operations research art to categorize items being analyzed in order to expedite processing and analysis of future products in the same category. Also, the examiner takes Official notice that minimization of end-of life inventory, i.e., inventory left over that must be written off, is a consideration in the field of product manufacturing.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Mujtaba to include Monte Carlo simulation after categorization of products by characteristics in order to predict performance of future products in the same category and to simulate the demand over the life of a product in order to minimize end-of-life inventory and the associated write-off costs.

Regarding claims 5 and 8, the examiner takes Official notice that a ratio is merely a comparison of two things and such a comparison is within the level of ordinary skill in the operations research field of endeavor.

Claims 12-21

The system claims, are unpatentable in view of Mujtaba, for the same reasons as set forth in the rejection above.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tamara L. Graysay whose telephone number is (571) 272-6728. The examiner can normally be reached on Mon - Fri from 8:30am to 5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tariq Hafiz, can be reached on (571) 272-6729. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Tamara L. Graysay
Examiner
Art Unit 3623

20050625

OIPE JC41
NOV 05 2001
PATENT & TRADEMARK OFFICE

Attachment 3, page 1 of 2

Sheet 1 of 1

FORM PTO-1449	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTY. DOCKET NO. 107101-00036
		SERIAL NO. 09/929,356
LIST OF REFERENCES CITED BY APPLICANT		APPLICANT YANAGINO et al.
(Use several sheets if necessary)		FILING DATE August 15, 2001
		GROUP 2163

U.S. PATENT DOCUMENTS

EXAMINER INITIAL		DOCUMENT NO.	DATE	NAME	CLASS	SUB-CLASS	FILING DATE
	AA						
	AB						
	AC						
	AD						
	AE						
	AF						

RECEIVED
NOV 07 2001
Technology Center 2100

FOREIGN PATENT DOCUMENTS

		DOCUMENT NO.	DATE	COUNTRY	CLASS	SUB-CLASS	TRANSLATION YES	NO	PART.
[initials]	AG	11-007482	01/12/1999	Japan					yes
	AH								
	AI								
	AJ								
	AK								
	AL								

OTHER REFERENCES *(Including Author, Title, Date, Pertinent Pages, Etc.)*

	AM	
	AN	
	AO	

EXAMINER T. GRAYSAY	DATE CONSIDERED 6/25/2005

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

TECH/48814.1

Part of Paper No 20050625

Attachment 3, page 2 of 2

FORM PTO-1449	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY DOCKET NUMBER 107101-00036	SERIAL NUMBER 09/929,356
LIST OF REFERENCES CITED BY APPLICANT (Use several sheets if necessary)		APPLICANT Tameo YANAGINO et al.	
		FILING DATE August 15, 2001	GROUP ART UNIT 2163

OIPE JC25 JUN 29 2004 PATENT & TRADEMARK OFFICE

U.S. PATENT DOCUMENTS

EXAMINER INITIAL		DOCUMENT NUMBER	DATE	NAME	CLASS	SUB-CLASS	FILING DATE
[initials]	AA	5,765,143	June 9, 1998	David E. SHELDON et al.			March 10, 1995
[initials]	AB	6,006,196	December 21, 1999	Gerald E. FEIGIN et al.			May 1, 1997
	AC						
	AD						
	AE						
	AF						

FOREIGN PATENT DOCUMENTS

		DOCUMENT NUMBER	DATE	COUNTRY	CLASS	SUB-CLASS	TRANSLATION YES	NO	PART.
	AG								
	AH								
	AI								
	AJ								
	AK								
	AL								

OTHER REFERENCES *(Including Author, Title, Date, Pertinent Pages, Etc.)*

[initials]	AM	Canadian Office Action dated March 11, 2004, Application Number: 2,381,569, Owner: Honda Giken Kogyo Kabushiki Kaisha, Title: METHOD OF FORECASTING FUTURE ORDERS IN PARTS INVENTORY MANAGEMENT, Classification: G06F-17/60
	AN	
	AO	

EXAMINER	DATE CONSIDERED
T. GRAYSAY	6/25/2005

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

TECH/248725.1

Part of Paper No 20050625

Attachment 1

Notice of References Cited	Application/Control No. 09/929,356	Applicant(s)/Patent Under Reexamination YANAGINO ET AL.	
	Examiner Tamara L. Graysay	Art Unit 3623	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-6,816,839	11-2004	Gung et al.	705/10
	B	US-6,611,726	08-2003	Crosswhite, Carl E.	705/28
	C	US-6,253,187	06-2001	Fox, Billy Shane	705/10
	D	US-6,144,945	11-2000	Garg et al.	705/28
	E	US-4,414,629	11-1983	Waite, John H.	705/28
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N	JP-08147357-A	06-1996	Japan	Sato	G06F 17/60
	O	JP-2000003388-A	01-2000	Japan	Arichika	G06F 17/60
	P	DE-19848094-A1	04-2000	Germany	Roennebeck	G06F 17/60
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Mujtaba, Enterprise modeling and simulation: complex dynamic behavior of a simple model of manufacturing, DEC1994, Hewlett-Packard Journal, v.45, n.6, p.80--113 (34 pages)
	V	Masters, Determinaton of near optimal stock levels for multi-echelon distribution inventories, !((#, Journal of Business Logistics, v.14, n.2, p.165-195 (31 pages)
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

N

PAT-NO: JP408147357A

DOCUMENT-IDENTIFIER: JP 08147357 A

TITLE: SIMPLY MODELING METHOD FOR MANUFACTURING DEVICE

PUBN-DATE: June 7, 1996

INVENTOR-INFORMATION:
NAME
SATO, AKIRA

ASSIGNEE-INFORMATION:

NAME	COUNTRY
NEC YAMAGATA LTD	N/A

APPL-NO: JP06287736

APPL-DATE: November 22, 1994

INT-CL (IPC): G06F017/60

ABSTRACT:

PURPOSE: To improve manufacture efficiency and to satisfy the demand of a customer by selecting the smallest value in the range of variation obtained by the data adoption rate and making the regression line at the time the characteristic model of a manufacture device.

CONSTITUTION: The part data on a lot size and a processing time is collected by a fixed period from each device. Then, regression analysis is executed between the average value of the processing time obtained by the lot size and the lot size to obtain the regression line so that the separation degree of the average value B4 from a value B3 on the regression line is obtd. in terms of ratio for each lot size. A value obtained by adding a value B5 most separated above from the regression line among the ratio values of a separation degree by the lot size and a value B6 most separated below from the regression line is used to define the range of variation. Next, a similar processing is executed by adopting the data of lower ratio. From among the ranges of variation obtained by adopting the data by ratio like this, the smallest value is selected and the regression line at the time is set as the characteristic model of the manufacture device.

COPYRIGHT: (C)1996,JPO

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号

特開平8－147357

(43)公開日　平成８年(1996)６月７日

(51)Int.Cl.⁶	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ　17/60				
			Ｇ０６Ｆ　15/ 21	Ｒ

審査請求　有　請求項の数１　ＯＬ　（全 ５ 頁）

(21)出願番号　特願平6－287736

(22)出願日　平成６年(1994)11月22日

(71)出願人　390001915
山形日本電気株式会社
山形県山形市北町４丁目12番12号

(72)発明者　佐藤　晃
山形県山形市北町四丁目12番12号　山形日本電気株式会社内

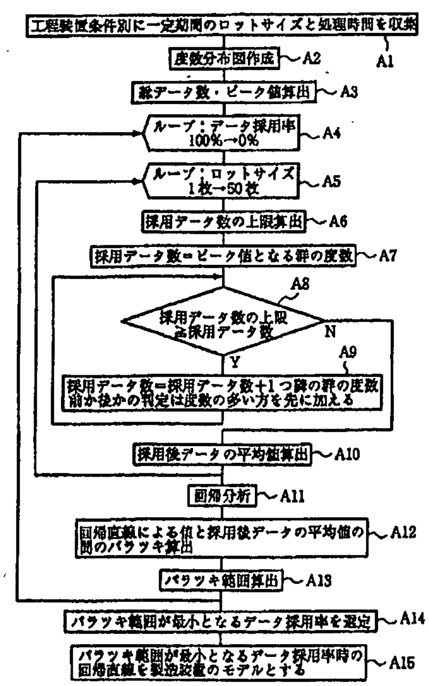
(74)代理人　弁理士　京本　直樹　（外２名）

(54)【発明の名称】　製造装置の簡易モデリング方法

(57)【要約】

【目的】より現実に近い製造装置の特性を簡易に抽出し、それを用いて製造装置の有効かつ効率的な利用を促すことにより生産効率を向上させ、かつ顧客の要求を満足させる（リードタイムの短縮）モデリング方法を提供する。

【構成】製品の処理時間が周期的に変化したり、外乱によって変化したり、一度に処理する製品のロットサイズの違いによって変化する製造装置のモデリングにおいて、製造装置の過去の一定期間における製品のロットサイズと製品の処理時間の関係からロットサイズ毎の処理時間に対してデータ採用率を変動させ、ロットサイズ毎に採用した処理時間の平均値から回帰分析を行ない、バラツキの範囲を導き出し、データ採用率毎に求めたバラツキの範囲の中で最も小さくなるデータ採用率を求め、その際の分析結果のデータを用いて各製造装置の現状の特性を容易に表現する。



【特許請求の範囲】

【請求項１】　製品の処理時間が周期的に変化したり、外乱によって変化したり、一度に処理する製品のロットサイズの違いによって変化する製造装置のモデリングにおいて、前記製造装置の過去の一定期間における製品のロットサイズと製品の処理時間の関係からロットサイズ毎の処理時間に対してデータ採用率を変動させ、ロットサイズ毎に採用した処理時間の平均値から回帰分析を行ない、バラツキの範囲を導き出し、データ採用率毎に求めたバラツキの範囲の中で最も小さくなるデータ採用率を求め、その際の分析結果のデータを用いて各製造装置の現状の特性を容易に表現できる機能を有することを特徴とする製造装置の簡易モデリング方法。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、製品の処理時間が周期的に変化したり、外乱によって変化したり、一度に処理するロットサイズの違いによって変化する製造装置を使用する生産ラインでの製造装置のモデリングにおいて、現在の製造装置の特性を簡易に把握できる、製造装置のモデリング方法に関する。

【０００２】

【従来の技術】以下に従来の技術について半導体ウェーハ生産ラインの一部の製造装置を例にして説明する。

【０００３】半導体ウェーハ生産ラインは、工程数が３００～５００工程と非常に多く、加えて、多品種少量生産に進み、全ての工程で製品処理条件の設定が多彩で頻繁に行われている。しかし、製造装置自体の設定は作業者の勘と経験に頼る面も見られ、自動化が遅れており、それが製品の処理時間に大きなバラツキを生じさせている。半導体ウェーハ生産ラインでは、このような製品の処理時間のバラツキの大きい製造装置を数多く使用しているため、全ての工程の製品装置を関連付けてうまく管理しなければ効率の良い生産ができない状況にある。

【０００４】近年、その管理方法として各製造装置をモデル化し、シミュレーションや処理能力計算を行ない、その結果を基に生産指示を出すというような方法が一般に使われ始めた。

【０００５】そのモデリング方法としては、あらゆる製造装置の詳細な動きを莫大なパラメータとして取り入れ、シミュレータや計算機の中で模擬的に動かすというものであり、莫大な工数を要し、製造装置の特性に変化がある度に製造装置の詳細な動きをパラメータとして取り入れなければならず、素早い対応が取れない。

【０００６】そこで、簡易的なモデリング方法として製造装置の動きを段取りと実作業に分け、それに要する時間をパラメータとするものが多く採用されている。

【０００７】しかし、上記したように製造装置自体の作業開始条件設定を作業者の勘と経験に頼っている部分が多いため、製造装置の処理時間のバラツキが大きく、履歴をとっても処理時間が周期的に変化したり、外乱によって変化したり、一度に処理するロットサイズの違いによって変化するため、正しい製造装置の特性が得られず、いかにその履歴からノイズデータを消した形で管理するかに焦点が当てられ、統計的手法による様々な試みがなされてきたが、高い精度でノイズを消すことのできる方法が確立されていなかったため現状に即した製造装置のモデリングができず、シミューレーションした結果と現実に大きな差を生じさせていた。

【０００８】

【発明が解決しようとする課題】前記の従来の製造装置のモデリング方法では、以下に示すような問題点があった。

【０００９】Ａ：頻繁に変化する製造装置の特性を取り入れることができない。

【００１０】Ｂ：頻繁に変化する製造装置の特性を取り入れるのに時間と工数を要する。

【００１１】Ｃ：オンライン化によるデータ精度の向上を図るためには莫大な費用を要する。

【００１２】Ｄ：現状の製造装置の特性をモデル化することが困難なため、シミュレーションや処理能力計算の結果と実際に差を生じさせる。

【００１３】これらの問題はライン全体の生産能力把握の精度を低下させるだけでなく、計画通りの生産を妨げる原因となり、会社全体の競争力を低下させ、かつ、経営状態を悪化させてしまうため、早急に改善する必要があった。

【００１４】本発明の目的は、より現実に近い製造装置の特性を簡易に抽出し、それを用いて製造装置の有効かつ効率的な利用を促すことにより生産効率を向上させ、かつ顧客の要求を満足させること（リードタイムの短縮）にある。

【００１５】

【課題を解決するための手段】本発明の製造装置の簡易モデリング方法は、まず、過去に処理したロットサイズと処理時間のデータを一定期間分集める。次に、ロットサイズ毎に処理時間別度数分布図を作成し、総データ数と度数がピークとなる群を求め、その場合の採用データ数の上限を決める。その際、ピーク値となる群の度数を採用データ数の初期値とする。

【００１６】次に、採用データ数が採用データ数の上限を超えていないかを比較し、超えていなければ採用データ数に度数分布図上で左右に隣接する群の内、度数の多い方を加え、再度採用データ数が採用データ数の上限を超えていないかを比較する。これを採用データ数の上限を超えるまで繰り返す。

【００１７】結果として採用されたデータの平均値を算出する。以上の処理をロットサイズ別に行う。

【００１８】次に、ロットサイズ別に求めた処理時間の平均値とロットサイズの間で回帰分析を行ない、回帰直

線を求める。ロットサイズ毎に回帰直線上の値に対するロットサイズ別に求めた処理時間の平均値の離れ具合を比率で求め、ロットサイズ毎の離れ具合比率値の内で回帰直線の上部に最も離れている値と、回帰直線の下部に最も離れている値を加えた値をバラツキの範囲とする。

【0019】ここまでを1サイクルとして、100％から0％までのデータ採用率別に同様の処理をする。

【0020】最後に採用率別に求めたバラツキの範囲の中から、その値が最も小さくなる値を選定し、そのときの回帰直線を製造装置の特性モデルとする。

【0021】

【実施例】本発明の1実施例について図1のフローチャ＊ートに基づき、かつ図2乃至図9を例示して説明する。

【0022】まず、ステップA1にて各装置から過去のロットサイズと処理時間のデータを一定期間分集める（例えば図2）。

【0023】次に、ステップA2にて、ロットサイズ毎に処理時間別度数分布図を作成（例えば図3）し、ステップA3にて、総データ数と度数がピークとなる群B2を求め（例えば図4）、ステップA4にてデータの採用率を100％とし、ステップA5にてロットサイズを1枚にする。ステップA6にてその場合の採用データ数の上限を決める。

【0024】

採用データ数の上限＝総データ数×データ採用率　…第1式

ステップA7にてピーク値となる群B2の度数を採用データ数B1の初期値とする。※

※【0025】

採用データ数B1＝ピーク値となる群B2の度数　…第2式

ステップA8にて採用データ数B1が採用データ数の上限を超えていないかを比較する（例えば図5）。

【0026】もし、超えていなければステップA9にて採用データ数B1に度数分布図上で左右に隣接する群の内、度数の多い方を加え（例えば図5）、ステップA8へと戻る。

【0027】これを超えるまで繰り返し、超えた時点（例えば図6）で、このループを抜けて、ステップA10へと進む。ステップA10では結果として採用されたデータの平均値を算出する。

【0028】以上の処理をロットサイズが50枚になるまで繰り返す。

【0029】次に、ステップA11にてロットサイズ別に求められた処理時間の平均値とロットサイズの間で回帰分析を行い、回帰直線を求める（例えば図7）。ステップA12にてロットサイズ毎に回帰直線上の値B3に対する平均値B4の離れ具合を比率で求める。

【0030】

離れ具合比率値＝（B4÷B3）－1　…第3式

ステップA13にてステップA12で求めたロットサイズ毎の離れ具合比率値の内で回帰直線の上部に最も離れている値B5と、回帰直線の下部に最も離れている値B6を加えた値をバラツキの範囲とする。

【0031】

バラツキの範囲＝｜B5｜＋｜B6｜　…第4式

ここまでを1サイクルとしてステップA4に戻り、次にデータ採用率を下げて同様に処理する。

【0032】例えばデータ採用率90％の場合で考えると、

前記第1式から、採用データ数の上限＝50件×90％＝45件

前記第2式から、採用データ数の初期値＝13件

となり、例えば図8のような採用となる。これをデータ採用率0％になるまで繰り返す。★

★【0033】ステップA14にてデータ採用率別に求めたバラツキの範囲の中から、その値が最も小さくなる値を選定する。ステップA15にてそのときの回帰直線を製造装置の特性モデルとする（例えば図9）。

【0034】

【発明の効果】以上に説明したように、本発明により、以下の効果が得られる。

【0035】（イ）：頻繁に変化する製造装置の特性を過去のデータから簡易にモデル化できる。

【0036】（ロ）：頻繁に変化する製造装置の特性を取り入れるための時間と工数を削減できる。

【0037】（ハ）：マニュアル作業の履歴からも製造装置の特性をモデル化できるため、オンライン化を最小限に抑えることができ、費用を削減できる。

【0038】（ニ）：現状の製造装置の特性をモデル化できるため、シミュレーションや処理能力計算の精度を向上できる。

【0039】以上により、会社全体の競争力が向上し、かつ経営状態の改善を図ることができる。

【図面の簡単な説明】

【図1】製造装置のモデリング方法の一例を示すフローチャートである。

【図2】ロットサイズ毎にデータをまとめた一例を示す図である。

【図3】ロットサイズ毎にまとめた度数分布の一例を示す図である。

【図4】ロットサイズ1枚・データ採用率100％時のピーク値の度数と上限の比較結果の例を示す図である。

【図5】ロットサイズ1枚・データ採用率100％時のピーク値の度数＋隣接する群の度数と上限の比較結果の例を示す図である。

【図6】ロットサイズ1枚・データ採用率100％時の最終採用データの例を示す図である。

【図7】データ採用率100％時の各ロットサイズ毎の

採用データの平均値による回帰直線の例を示す図である。

【図8】ロットサイズ１枚・データ採用率９０％時の最終採用データの例を示す図である。

【図9】バラツキの範囲が最小となるデータ採用率時の回帰直線の例を示す図である。

【符号の説明】

Ａ１～Ａ１５　各ステップ

Ｂ１　採用データ数

Ｂ２　度数がピークとなる群

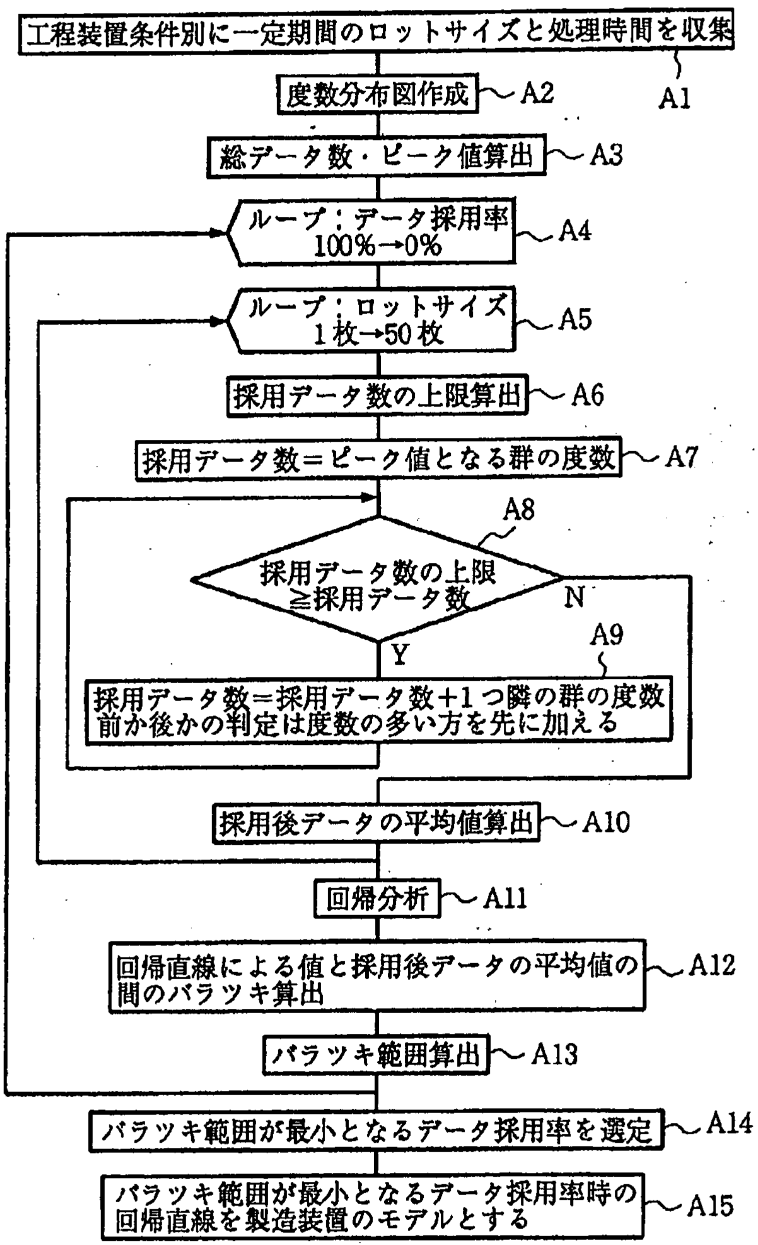
Ｂ３　回帰直線上の値

Ｂ４　採用データの平均値

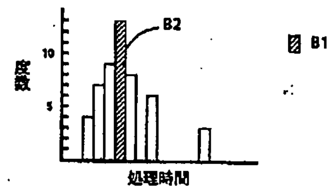
Ｂ５　回帰直線の上に最も離れている値

Ｂ６　回帰直線の下に最も離れている値

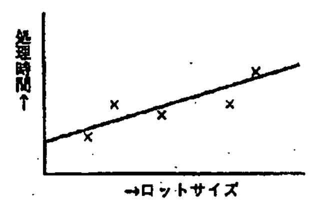
【図1】



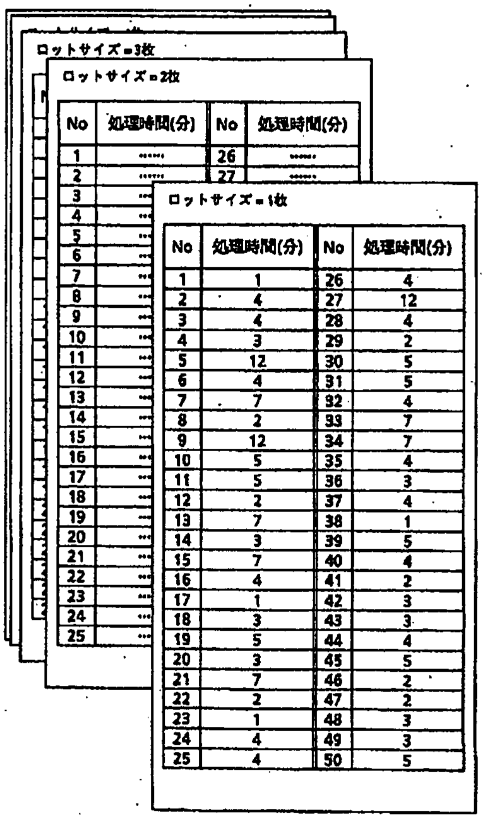
【図4】



【図9】

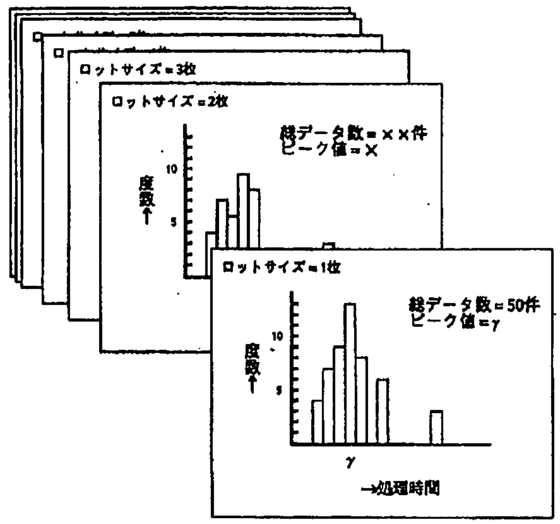


【図2】

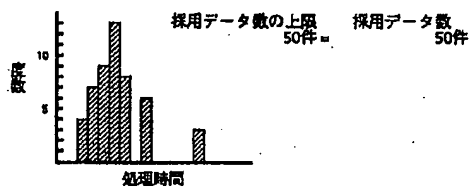


No	処理時間(分)	No	処理時間(分)
1	1	26	4
2	4	27	12
3	4	28	4
4	3	29	2
5	12	30	5
6	4	31	5
7	7	32	4
8	2	33	7
9	12	34	7
10	5	35	4
11	5	36	3
12	2	37	4
13	7	38	1
14	3	39	5
15	7	40	4
16	4	41	2
17	1	42	3
18	3	43	3
19	5	44	4
20	3	45	5
21	7	46	2
22	2	47	2
23	1	48	3
24	4	49	3
25	4	50	5

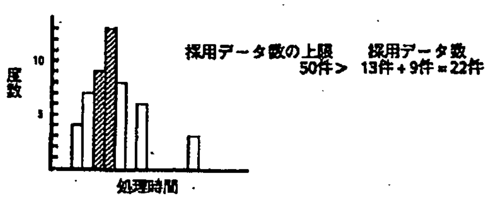
【図3】



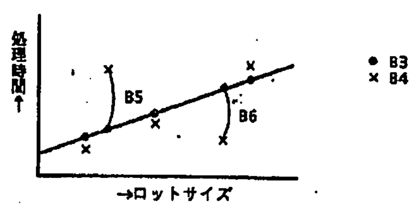
【図6】



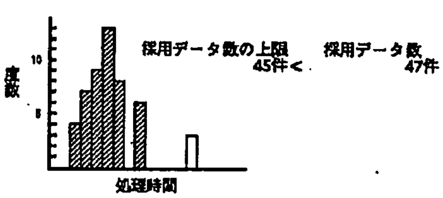
【図5】



【図7】



【図8】



0

PAT-NO: JP02000003388A

DOCUMENT-IDENTIFIER: JP 2000003388 A

TITLE: DEMAND FORECASTING UNIT SETTING METHOD

PUBN-DATE: January 7, 2000

INVENTOR-INFORMATION:

NAME	COUNTRY
ARICHIKA, SUSUMU	N/A
SATO, TAKAO	N/A

ASSIGNEE-INFORMATION:

NAME	COUNTRY
HITACHI LTD	N/A

APPL-NO: JP10168028

APPL-DATE: June 16, 1998

INT-CL (IPC): G06F017/60

ABSTRACT:

PROBLEM TO BE SOLVED: To dynamically extract a suitable category and to grasp the most approximate demand trend by selecting a category capable of recognizing the demand trend and deciding demand forecasting quantity based on the demand result of the selected category.

SOLUTION: A forecasting object commodity name whose demand is to be forecast is obtained by an input device 114, and the factor and weighting of the forecasting object commodity are obtained from a commodity factor DB 100. The taken out factor whose content by individual commodity is equal to that of the forecasting object commodity is selected and is extracted as one category in a category classification DB 102. The real demand for an arbitrary period of the commodity belonging to the extracted category is obtained from a real demand DB 104 by individual products, they are time-sequentially summed up and the real demand of the category is calculated. Multiple regression analysis is executed on the calculated real demand and it is judged whether the demand trend is recognized or not. When the demand trend can be recognized, the demand forecasting quantity at the demand forecasting object date of the category is calculated by regression analysis and demand forecasting quantity is decided.

COPYRIGHT: (C)2000,JPO

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号
特開2000－3388
（P2000－3388A）
(43)公開日　平成12年1月7日(2000.1.7)

(51)Int.Cl.⁷	識別記号	ＦＩ	テーマコード（参考）
Ｇ０６Ｆ　17/60		Ｇ０６Ｆ　15/21	Ｚ　５Ｂ０４９

審査請求　未請求　請求項の数４　ＯＬ　（全 5 頁）

(21)出願番号　特願平10－168028

(22)出願日　平成10年6月16日(1998.6.16)

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72)発明者　有近　晋
神奈川県横浜市都筑区加賀原二丁目2番
株式会社日立製作所システム開発本部内

(72)発明者　佐藤　隆夫
神奈川県横浜市都筑区加賀原二丁目2番
株式会社日立製作所システム開発本部内

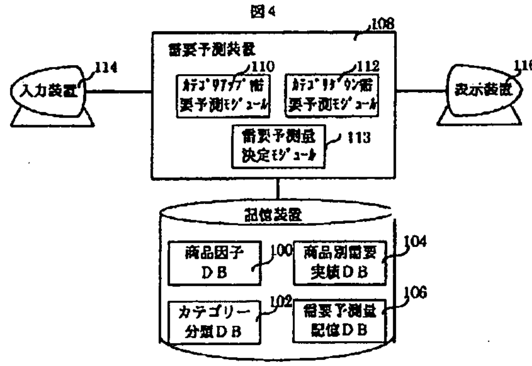
(74)代理人　100068504
弁理士　小川　勝男

Ｆターム(参考)　5B049 BB13 CC00 CC31 EE03 FF01
FF07

(54)【発明の名称】　需要予測単位設定方法

(57)【要約】
【課題】本発明は、需要予測を行うのに好適なカテゴリーを動的に抽出することにより需要予測を行う方法を提供することにある。
【解決手段】因子と需要実績を用いて需要予測を行う方法において、使用する因子を組み合わせることにより、需要動向が認められるぎりぎりのカテゴリーを抽出し、（ステップ２６及びステップ４８）、抽出されたカテゴリーと予測対象の需要実績の比率から需要予測量を算出する（ステップ１０）。

図４



【特許請求の範囲】

【請求項１】商品または市場の需要動向が認められない場合、需要動向に影響を与える要素を選択し、需要動向に傾向が見られるような、需要動向に影響を与える要素の内容が同一の商品の集まりを抽出し、それを基にして需要予測を行うことを特徴とする需要予測単位設定方法。

【請求項２】請求項１において、需要動向に影響を与える要素の内容が同一の商品の集まりを抽出する際、需要動向に影響を与える要素の重要度に関する情報から重要度の小さいものから順に選択する対象から外していくという方法と、需要動向に影響を与える要素の重要度に関する情報から重要度の大きいものから順に選択する対象に加えていくという方法という２つの方法を用いて、抽出対象となっている商品数がなるべく少ない需要動向に影響を与える要素の内容が同一の商品の集まりを抽出し、その需要実績に関する情報を基にして需要予測を行った後、信頼性が高いものを選択することを特徴とする需要予測単位設定方法。

【請求項３】請求項１において、需要動向に影響を与える要素の重要度に関する情報から重要度の小さいものから順に選択する対象から外していくという方法のみを用いて需要動向に影響を与える要素の内容が同一の商品の集まりの抽出を行い、その需要実績に関する情報を基にして需要予測を行うことを特徴とする需要予測単位設定方法。

【請求項４】請求項１において、需要動向に影響を与える要素の重要度に関する情報から重要度の大きいものから順に選択する対象に加えていくという方法のみを用いて需要動向に影響を与える要素が同一の商品の集まりの抽出を行い、その需要実績に関する情報を基にして需要予測を行うことを特徴とする需要予測単位設定方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は需要予測対象商品または市場の需要動向を認識できない際、使用する需要動向に影響を与える要素を選択しつつ、需要動向が認められる需要動向に影響を与える要素が同一の商品の集まりを抽出して、需要予測量決定に活用する方法に関する。

【０００２】

【従来の技術】従来は、需要予測を行うための、需要動向に影響を与える要素（以下、因子と略）が同一の商品及び市場（以下、カテゴリーと略）は事前に設定しておくものであった。例えば、特開平７−３６８５４号公報に示されるように、特性を反映させるために、予測対象商品を複数のカテゴリーを利用して予測する場合でも、予測対象商品がどのカテゴリーに所属するかは事前に決定されていた。この技術では、データベースのデータを広域のマクロデータ、小域のミクロデータ、及び業界データの複数グループに分け、該データの項目を説明変数とし、重回帰分析により予測データを求めることができる。

【０００３】

【発明が解決しようとする課題】近年の商品品種の増加及び、商品ライフサイクルの短期化に伴い、実績情報が十分に揃えられない状況が発生している。また、その一方で在庫圧縮のため需要予測の精度向上に対する要望は高い。この結果、次のような問題が発生した。

【０００４】商品が市場投入間もなく、需要動向を促えるだけの十分な需要実績が蓄積されていない。

【０００５】商品の需要量が希少であり、需要動向を的確に捉えられない。

【０００６】商品の需要動向に与える因子が異なり、需要動向が的確に捉えられない。

【０００７】その結果、従来は需要実績が十分でないため、担当者のカンによって需要予測量を決定したり、不適当なカテゴリーから需要予測量を算出するため、適切な需要予測量を算出することが困難であるという状況が発生していた。

【０００８】本発明の目的は、需要動向が的確に捉えられない場合、好適なカテゴリーを動的に抽出することにより、極近値の需要動向を捉える方法を提供することにある。

【０００９】

【課題を解決するための手段】本発明は需要予測において、因子に関する情報より使用する因子を選択し、選択された因子と同じ因子の内容を持つカテゴリーをカテゴリー分類に関する情報より選別するという方法によりカテゴリーを抽出し、需要傾向が認められるカテゴリーを選択し、選択されたカテゴリーの需要実績を元に需要予測量を決定する。例えば、予測対象商品または市場の持つ因子から使用する因子を選定し、カテゴリー分類に関する情報から、因子の内容が同一の商品を選択し、カテゴリーを形成する。形成されたカテゴリーに属する各商品の需要実績情報からカテゴリーの需要実績を算出し、この値が需要動向が認められるものであった場合、その値を元にして予測対象商品または市場の需要予測値を算出する。需要動向が安定していなかった場合、使用する因子を見直し、再度カテゴリーを形成する。

【００１０】

【発明の実施の形態】本発明は、因子を組み合わせることにより、需要動向が安定するぎりぎりのカテゴリーを動的に抽出し、それに基づき需要予測を行う方法である。

【００１１】以下、本発明の実施の形態を詳細に説明する。本発明は商品の需要予測のみならず、市場の需要予測にも適用可能であるが、商品の需要予測を行う際の実施の形態を例に取り説明する。市場に関しても同様の方法で実施可能である。

【００１２】図1は、本発明の全体像を示すＰＡＤ図で

ある。図2は図1のステップ1を詳細化したＰＡＤ図である。図3は図1のステップ5を詳細化したＰＡＤ図である。図4は本発明の構成を示したブロック図である。図5は商品因子ＤＢが持つべきデータテーブルである。図6はカテゴリー分類ＤＢが持つべきデータテーブルである。図7は商品別需要実績ＤＢが持つべきデータテーブルである。図8は需要予測量決定のモデル図である。

【００１３】図4において商品因子ＤＢ１００は商品別の因子２０２及びその重みづけ２０２を登録している。カテゴリー分類ＤＢ１０２は各商品の因子毎の特性２０４を登録している。商品別需要実績ＤＢ１０４は商品別の過去の需要実績２０６を登録している。需要予測量記憶装置１０６は算出された需要予測量を登録する。需要予測装置１０８は商品因子ＤＢ１００やカテゴリー分類ＤＢ１０２や商品別需要実績ＤＢ１０４の情報を受取り、その情報を用いて需要予測処理を行う。需要予測処理は、カテゴリーアップ需要予測モジュール１１０及びカテゴリーダウン需要予測モジュール１１２を用いて行う（カテゴリーアップ及びカテゴリーダウンは後に説明）。需要予測量決定モジュール１１３は最終的な需要予測量を決定する。入力装置１１４は需要予測を指示する情報の入力を行い、表示装置１１６は需要予測を行う際の有意情報の提供を行う。

【００１４】次に、図1，図2及び図3のＰＡＤ図に基づいて図４各部の動作を説明する。

【００１５】まず、需要予測を行うべき予測対象商品名を入力装置１１４より入手する。この後、需要動向に影響を与える要素の重要度に関する情報から重要度の小さいものから順に選択する対象から外していくという方法（以下、カテゴリーアップと略）を用いて需要予測を行う（ステップ1）。入手された予測対象商品の因子２００及びその重みづけ２０２を商品因子ＤＢ１００より入手する。因子数量指定装置（図に表示せず）により指定された任意の数量の因子を重みづけ２０２の上位のものを優先して取り出す（ステップ４０）。カテゴリー分類ＤＢ１０２において、取り出された因子の商品別の内容２０４が予測対象商品と同一のものを選択し、一つのカテゴリーとして抽出する（ステップ４２）。

【００１６】抽出されたカテゴリーに属する商品の任意の期間の需要実績を商品別需要実績ＤＢ１０４より入手し、時系列的に合計してカテゴリーの需要実績を算出する（ステップ２４）。算出された需要実績に対し重回帰分析を行い、危険率記憶装置（図示せず）より入手された任意の危険率とｆ分布記憶装置（図示せず）より入手されたｆ分布より算出される値よりも、分散比（回帰による変動の不遍分散／残差の変動の不偏分散により算出）が大きければ需要動向が認められる、小さければ需要動向が認められないと判断する（ステップ２６）。

【００１７】需要傾向が認められなかった場合、取り出されている因子の重みづけ２０２の下位のものを一つ対

象から外す（ステップ２８）。これはカテゴリーの対象となる商品の数量を増やすことによって需要動向がみとめられ易くするための処理である。需要動向が認められるようになるまで、ステップ２２～ステップ２８を繰り返す。需要動向が認められたら、カテゴリーの需要予測対象日時の需要予測量を回帰分析により算出し、その値とカテゴリー全体の需要実績と予測対象商品単独の需要量の比率から需要予測量を決定する（ステップ３０）。ステップ３０を概念的に示したのが図8である。

【００１８】次に、需要動向に影響を与える要素の重要度に関する情報から重要度の大きいものから順に選択する対象に加えていくという方法（以下、カテゴリーダウンと略）を用いて需要予測を実施する（ステップ2）。因子数量指定装置（図に表示せず）により指定された任意の数量分だけ予測対象商品の因子２００の重みづけ２０２が上位のものから取出す（ステップ４０）。ステップ２２と同じ方法によってカテゴリーを抽出する（ステップ４２）。以前に算出したカテゴリーの需要実績があるならば、それを一時的にメモリー上（図示せず）に保存する。

【００１９】今回抽出されているカテゴリーの需要実績をステップ２４と同じ方法により算出する（ステップ４６）。ステップ２６と同じ方法により需要動向安定チェックを行う（ステップ４８）。需要動向が認められた場合、現在使用されていない因子２００の中で重みづけ２０２が上位の因子２００を取り出す（ステップ５０）。その後、初めて需要動向が認められないカテゴリーが出るまでステップ４２～５０を繰り返す。需要動向が認められないカテゴリーが出たら、旧需要実績を基に需要予測対象日時の需要予測量を回帰分析により算出し、旧需要実績と予測対象商品の需要実績の比率により需要予測量を算出する（ステップ５２）。

【００２０】次に、カテゴリーアップ時の最後のカテゴリーの商品数とカテゴリーダウン時の旧需要実績を算出した時のカテゴリーの商品数を比較し、商品数が少ない方が算出した需要予測量を最終的な需要予測量として決定し、需要予測量記憶ＤＢ１０６に保存する（ステップ１０）。

【００２１】またカテゴリーアップ需要予測のみを行って、その結果を需要予測量としても良い。

【００２２】またカテゴリーダウン需要予測のみを行って、その結果を需要予測量としても良い。

【００２３】また、ステップ４０において、因子数量指定装置（図に表示せず）により指定された任意の数量ではなく、入力装置１１４により随時、任意の数量を指定でも良い。

【００２４】また、ステップ２６及びステップ４８において需要動向が認められるかどうかの判断を、予測対象商品の需要実績、カテゴリーの需要実績、分散等を表示装置１１６に表示し、需要動向がみとめられるかどうか

の判断を入力装置１１４によりおこなっても良い。

【００２５】また、季節変動を考慮した需要予測も統計的に処理することにより可能である。

【００２６】また、ステップ１０において、商品数、分散等を表示装置１１６に表示し、どちらを最終的な需要予測量とするかの判断を、入力装置１１４により行っても良い。

【００２７】また、ステップ２０において、予測対象商品の因子２００及びその重みづけ２０２を商品因子ＤＢ１００より入手するのではなく、因子２００とその重みづけ２０２を、入力装置１１４により入手することも可能である。

【００２８】

【発明の効果】以上に述べたように、本発明によれば、需要予測を実施する際カテゴリーを動的に設定できるため、需要動向が不安定であったり、実績データが十分でないと言った場合でも、好適なカテゴリーを抽出して需要予測を実施できる。

【図面の簡単な説明】

【図１】本発明の処理手順の概要を示すＰＡＤ図である。

【図２】本発明のカテゴリーアップの処理手順の実施の形態を示すＰＡＤ図である。

【図３】本発明のカテゴリーダウンの処理手順の実施の形態を示すＰＡＤ図である。

【図４】本発明に係わる需要予測装置のブロック図である。

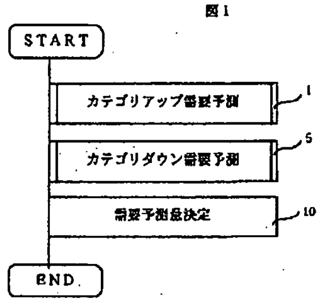
【図５】商品因子情報の構成例を示す図である。

【図６】カテゴリー分類情報の構成例を示す図である。

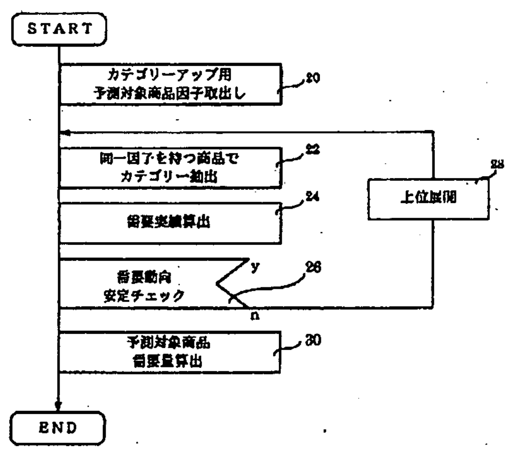
【図７】商品別需要動向情報の構成例を示す図である。

【図８】需要予測量の算出方法の実施例を示す特性図である。

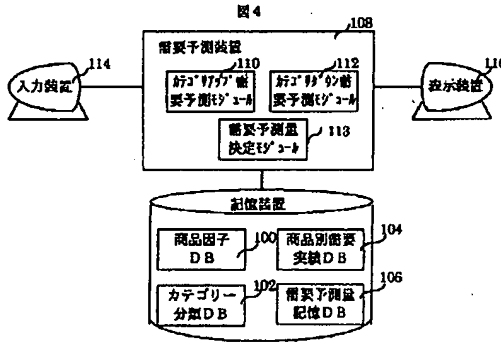
【図１】



【図２】



【図４】



【図５】

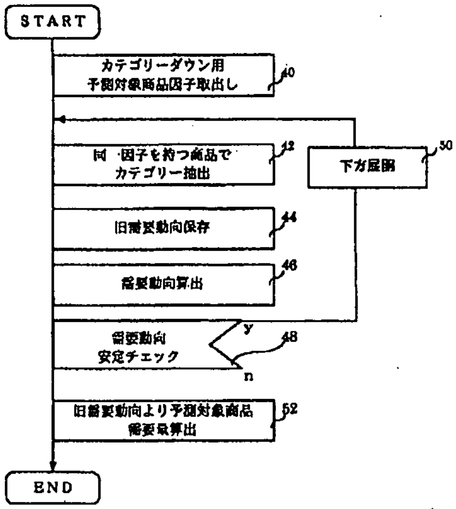
図５

商品因子ＤＢ　200　202

因子	重みづけ
顧客層	7
競合他社状況	9
パッケージ	5
材料	3
消費者必需度	7
プロモーション状況	5
店舗分布	5

【図3】

図3



【図6】

図6

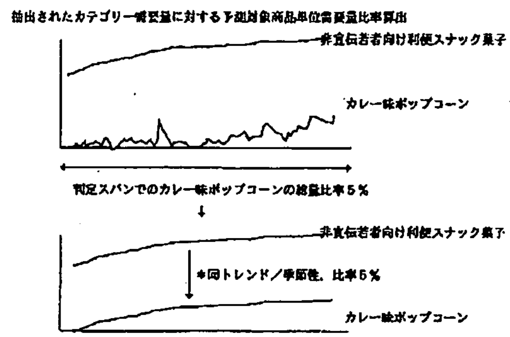
カテゴリー分類DB

204

	プロモーション状況	顧客層	商品ステータス	テイスト	・・・
カレー味ポップコーン	1	1	1	1	
チーズ味ポップコーン	1	2	2	1	
イカスミ味ポップコーン	3	3	3	1	
メロンチョコレート	2	1	4	2	
コーヒーキャンディ	1	1	1	2	

【図8】

図8



【図7】

図7

商品別需要実績DB

206

	1	2	3	4	5	6	7	8	9	10
カレー味ポップコーン	100	150	20	50	120	115	200	250	300	200
チーズ味ポップコーン	40	60	30	16	10	5	12	15	12	8
イカスミ味ポップコーン	60	62	75	23	21	18	12	8	23	35

P

DERWENT-ACC-NO: 2000-319316

DERWENT-WEEK: 200031

COPYRIGHT 2005 DERWENT INFORMATION LTD

TITLE: Determining degree to which tasks have been completed in technical systems by using Monte Carlo simulation to reproduce effects of components on whole system

INVENTOR: ROENNEBECK, H

PATENT-ASSIGNEE: ASEA BROWN BOVERI AG[ALLM]

PRIORITY-DATA: 1998DE-1048094 (October 19, 1998)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
DE 19848094 A1	April 20, 2000	N/A	008	G06F 017/60
EP 996044 A2	April 26, 2000	G	000	G05B 017/02

DESIGNATED-STATES: AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO	APPL-DATE
DE 19848094A1	N/A	1998DE-1048094	October 19, 1998
EP 996044A2	N/A	1999EP-0810877	September 29, 1999

INT-CL (IPC): G05B017/02, G06F017/50 , G06F017/60

ABSTRACTED-PUB-NO: DE 19848094A

BASIC-ABSTRACT:

NOVELTY - The method involves using a Monte Carlo simulation to represent the functional cooperation between individual components of the system as functional blocks. These functional blocks reproduce the effect of failure of one component on the whole system. Changes in the availability of the system due to random and planned events are weighted using factors formed from the degree of severity of non-availability. The effects of events such as repair and maintenance measures on the system availability are taken into account.

USE - E.g. for power generation plants, vehicles.

ADVANTAGE - Takes account of maintenance and parts replacement strategies.

DESCRIPTION OF DRAWING(S) - The drawing shows a flow diagram of the method.

CHOSEN-DRAWING: Dwg.3/3

TITLE-TERMS: DETERMINE DEGREE TASK COMPLETE TECHNICAL SYSTEM SIMULATE REPRODUCE EFFECT COMPONENT WHOLE SYSTEM

DERWENT-CLASS: T01

EPI-CODES: T01-J05A2; T01-J15H;

SECONDARY-ACC-NO:
Non-CPI Secondary Accession Numbers: N2000-239571

(19) BUNDESREPUBLIK DEUTSCHLAND

DEUTSCHES PATENT- UND MARKENAMT

(12) **Offenlegungsschrift**

(10) **DE 198 48 094 A 1**

(51) Int. Cl.⁷:
G 06 F 17/60
G 06 F 17/50

(21) Aktenzeichen: 198 48 094.6
(22) Anmeldetag: 19. 10. 1998
(43) Offenlegungstag: 20. 4. 2000

DE 198 48 094 A 1

(71) Anmelder:
Asea Brown Boveri AG, Baden, Aargau, CH

(74) Vertreter:
Lück, G., Dipl.-Ing. Dr.rer.nat., Pat.-Anw., 79761 Waldshut-Tiengen

(72) Erfinder:
Rönnebeck, Horst, Dr., 79780 Stühlingen, DE

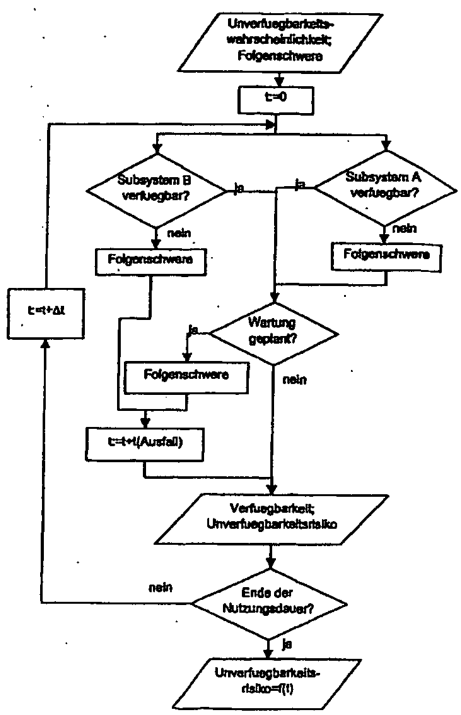
(56) Für die Beurteilung der Patentfähigkeit in Betracht zu ziehende Druckschriften:

DE 197 15 715 A1
DE 197 13 917 A1
DE 297 12 807 U1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Verfahren zur Ermittlung des Grades der Aufgabenerfüllung technischer Systeme

(57) Ein Verfahren zur Bestimmung des Grades der technischen Aufgabenerfüllung technischer Systeme berücksichtigt neben der reinen Verfügbarkeit des Systems auch die Folgenschwere eines Unverfügbarkeitsereignisses. In die Berechnung des so ermittelten Unverfügbarkeitsrisikos fließt als Folgenschwere der Unverfügbarkeit auch der Ressourcenverbrauch bei der Behebung derselben ein, ebenso wie die Möglichkeit, im Falle der Unverfügbarkeit Ersatzsysteme bereitzustellen. Beispielsweise kann zwischen geplanten und ungeplanten Unverfügbarkeiten unterschieden werden. Das Verfahren simuliert die Nutzungsdauer eines Systems als eine Folge von Durchläufen einer Monte-Carlo-Simulation, wobei innerhalb einer Simulation durch Funktionsblöcke die Interdependenz der einzelnen Komponenten nachgebildet wird. Die empirisch oder teilempirisch ermittelte Versagenswahrscheinlichkeit der Einzelkomponenten dient hierbei als Eingabegröße. Ereignisse, die während eines Simulationsdurchlaufes auftreten, werden in den folgenden Simulationen berücksichtigt. Somit kann das reale Verhalten des Systems mit einfachen Vorgaben sehr genau beschrieben werden. Mit Hilfe des Verfahrens kann die Auslegung und die Instandhaltung des Systems so gesteuert werden, dass ein maximaler Grad der Erfüllung der technischen Aufgabe bei minimalem Ressourceneinsatz resultiert.



DE 198 48 094 A 1

Beschreibung

Technisches Gebiet

Die vorliegende Erfindung betrifft ein Verfahren nach dem Oberbegriff des Anspruchs 1.

Stand der Technik

Vor dem Hintergrund einer ständig steigenden Weltbevölkerung gerade in den bis anhin schwächer entwickelten Regionen, der Notwendigkeit, auch diese Regionen zu industrialisieren, und dem daraus resultierenden Hunger nach natürlichen Resourcen, trat im Laufe der vergangenen beiden Jahrzehnte in der modernen Technik der effiziente und nachhaltige Einsatz von Rohstoffen und Energie zusehends in den Vordergrund.

Neben dem effizienten Resourceneinsatz beim Betrieb eines technischen Systems ist hierbei auch der Umstand zu sehen, dass technische Systeme selbst Resourcen binden, insbesondere in Form in ihnen verbauter Rohstoffe, und während ihrer Herstellung eingesetzter Energie, die sich trotz vermehrter Anstrengungen, den Kreislauf wirtschaftlich verwerteter Stoffe zu schliessen, meist nur unvollkommen und wiederum unter hohem neuerlichem Resourceneinsatz wiederverwerten lassen.

Die naheliegendste Möglichkeit, den Resourceneinsatz zu minimieren, besteht sicherlich in konsequentem Leichtbau und daraus resultierenden Materialeinsparungen, wie auch in vielen Fällen Energieeinsparungen. Jedoch ist zu berücksichtigen, dass ein technisches System nur dann überhaupt einen sinnvollen Resourceneinsatz darstellt, wenn es in der Lage ist, die gestellte Aufgabe mit hoher Effizienz zu verrichten. Daraus ergibt sich, dass ein System, das mit geringem Rohstoffeinsatz unter Verwendung modernster Technologien leicht und filigran konstruiert wurde, aufgrund einer verstärkten Belastung möglicherweise einen hohen Wartungsaufwand, höhere Ausfallrisiken und eine geringere Lebensdauer aufweist. Ab einer gewissen Grenze der Resourcenausnutzung in einem technischen System wird also der Fall auftreten, dass einzelne Komponenten häufig ausgetauscht werden müssen, dass häufige Stillstände erforderlich sind, und dass nicht zuletzt durch eine geringere Verfügbarkeit der Grad der Erfüllung der technischen Aufgabe sinkt, so, dass eine stärkere Redundanz des Systems erforderlich ist. In der Summe wird die ursprünglich erzielte Resourceneinsparung also wieder wettgemacht.

In diesem Zusammenhang darf auch nicht auf den reinen Verbrauch stofflicher Resourcen fokussiert werden: Ein hochkomplexes System höchster technischer Wertigkeit mit entsprechendem Wartungsbedarf bindet während seiner Planung, seiner Erstellung und seinem Einsatz hochqualifizierte Arbeitskraft, die bei der erforderlichen Spezialisierung ebenfalls zu einem knappen Gut wird.

Dagegen steht die Philosophie, technische Systeme sehr robust und einfach auszulegen. Aus dem Einsatz weniger, nur gering belasteter Komponenten resultiert ein zuverlässiges System mit geringem Wartungsaufwand und einem geringen Bedarf an Redundanz, das allerdings bei seiner Ausführung einen hohen Resourceneinsatz erfordert, und weniger effizient betrieben werden kann. Hinzu kommt, dass möglicherweise dieses System weit vor dem Ende seiner Lebensdauer veraltet, das heisst gegenüber dem Stand der Technik eine derart geringe Wertigkeit aufweist, dass es nicht länger in nennenswertem Ausmasse betrieben wird, woraus letztlich wieder eine mangelhafte Ausnutzung zusehends knapper und wertvoller Resourcen resultiert.

Der verantwortungsvolle Umgang mit natürlichen Resourcen stellt somit eine höchst komplexe Optimierungsaufgabe dar, deren Lösung davon abhängt, die Zusammenhänge richtig darzustellen und die unterschiedlichen Einflussgrössen möglichst realistisch zu bewerten.

Diese abstrakten Ausführungen seien nachfolgend anhand eines Beispiels dem interessierten Fachmann nahegebracht.

Es sei die Versorgung eines Netzes mit einer bestimmten elektrischen Leistung sicherzustellen. Einfach und schnell – auch mit geringem Rohstoffeinsatz – ist dies mit einer Anzahl Gasturbinen zu bewerkstelligen. Ausgangspunkt der Betrachtung seien Gasturbinen einer älteren bewährten Baureihe, mit moderaten Heissgastemperaturen und einem moderaten Brennkammerdruck. Hier kann mit einer hohen Zuverlässigkeit und geringem Wartungsaufwand, somit hoher Verfügbarkeit, gerechnet werden, allerdings auf Kosten des Wirkungsgrades. Insgesamt wird sich eine relativ geringe technische Wertigkeit des Systems ergeben, bedingt durch eine geringe Resourcenausnutzung, auch aufgrund geringer Materialauslastungen. Andererseits muss zur Aufrechterhaltung der technischen Funktion auch eine geringe Reserve vorgehalten werden, was den Resourceneinsatz verringert.

Hypothetisch werden nun die thermodynamischen Daten der Gasturbinen gesteigert. Damit steigt der Wirkungsgrad einer Maschine, und, da das Material höher belastet wird, auch die Rohstoffausnutzung. Im Gegenzug wird die Lebensdauer und Zuverlässigkeit bestimmter Komponenten, und der Wartungsaufwand, steigen. Ein höherer Wartungsaufwand bedeutet mehr Stillstandzeiten, weshalb mehr Reserve vorgehalten werden muss, anders ausgedrückt, es müssen Resourcen als Reserve für Unverfügbarkeiten vorgehalten werden, und diese Resourcen tragen nicht primär zur Funktion des technischen Systems bei. Das heisst, bei einer Wirkungsgraderhöhung der Gasturbinen steigt die Resourcenausnutzung zunächst, sinkt aber ab einem bestimmten Punkt aufgrund der notwendigen hohen Reservevorhaltung bezogen auf das Gesamtsystem "Stromversorgung" wieder.

Nunmehr kann bei einem bestimmten Entwicklungsstand der Gasturbinen ein Kombikraftwerk in Erwägung gezogen werden. Damit steigt der Rohstoffeinsatz bei der Herstellung, jedoch sinkt der Resourcenverbrauch bei der Stromproduktion. Andererseits steigt die Zahl der Komponenten, die einen Ausfall des Systems verursachen können, und es stellt sich wieder die Frage nach der notwendigen Redundanz zur Aufrechterhaltung der Funktion des Gesamtsystems.

Eine realistische Beurteilung der Verfügbarkeit eines technischen Systems, respektive der notwendigen Reservevorhaltung bedingt nunmehr, das Ausfallrisiko einer einzelnen Komponente zu kennen, und auch die Folgenschwere eines Ausfalls zu beurteilen: Nicht jede Komponente innerhalb eines Kraftwerks ist von so unmittelbarer Bedeutung für die technische Funktion, dass sein Ausfall mit dem Ausfall des Gesamtsystems gleichzusetzen ist. Mitunter kann eine Reparatur bis zur nächsten Wartung aufgeschoben werden. Auch ist die Folgenschwere einer Unverfügbarkeit des gesamten Systems nicht immer dieselbe: Eine geplante Unverfügbarkeit zu Wartungszwecken ist sicher leichter beherrschbar als ein plötzlicher Ausfall eines Kraftwerkes.

Hieraus resultiert weiterhin der Zwang, bei der Verfügbarkeitsanalyse die funktionalen Zusammenhänge zwischen den einzelnen Komponenten des Gesamtsystems, die und die Folgenschwere von Ausfällen zu berücksichtigen. Weiterhin gilt es auch, die Wartungs- und Teileaustauschstrategie in Betracht zu ziehen. Die Wahrscheinlichkeit einer unvermittelten Unverfügbarkeit sinkt sicher, wenn Wartungsmassnahmen rechtzeitig durchgeführt werden. Andererseits

ist auch die Wartung eine Unverfügbarkeit, und der Austausch von Teilen ein Resourcenverbrauch. Es kann ein Bauteil ausgetauscht worden sein, das noch lange gehalten hätte, und durch eines ersetzt worden sein, das durch einen Fabrikationsfehler recht bald versagt. Ebenso kann es bei der Wartung selbst zu Fehlern kommen, was wiederum durch den Verbrauch der Resource "Qualifiziertes Personal" beeinflusst wird.

Es wird deutlich, welche Faktoren bei ganzheitlichen Beurteilung der Zuverlässigkeit eines technischen Systems zu berücksichtigen sind: Neben der reinen Unverfügbarkeitswahrscheinlichkeit muss auch die Folgenschwere einer Unverfügbarkeit berücksichtigt werden, worunter die Nichterfüllung der technischen Aufgabe, der zur Behebung einer Unverfügbarkeit notwendige Resourcenverbrauch – Rohstoffe und qualifizierte Arbeitskraft- sowie ein zur Reservevorhaltung notwendiger Resourcenverbrauch zu subsummieren sind, wobei diese Aufzählung keinen Anspruch auf Vollständigkeit erhebt. Weiterhin ist sinnvollerweise der gesamte zur Bereitstellung des technischen Systems notwendige Resourceneinsatz zu gewichten: Ein vielfach redundantes System wird zu einer geringen Unverfügbarkeitswahrscheinlichkeit führen, wobei das System aber unverhältnismässig viele Resourcen bindet.

Schliesslich muss noch berücksichtigt werden, dass eine neue Technologie, die prima facie ein höheres Unverfügbarkeitsrisiko beinhaltet, über die Lebensdauer eines technischen Systems wie eines Kraftwerks weiterentwickelt wird, und die Verfügbarkeit durch sukzessiven Austausch von Komponenten gegen verbesserte Varianten erhöht wird.

Dieser unvollständige Abriss zeigt die enorme Komplexität des der Erfindung zugrundeliegenden Problems. Derartige Verfügbarkeitsanalysen werden bis anhin beispielsweise mit Hilfe der Fehlerbaumanalyse, der Ereignisbaumanalyse, oder der Markov-Analyse durchgeführt. Mit allen genannten Verfahren ist es schwierig, zu einer Beurteilung der Folgenschwere einer Unverfügbarkeit, zu kommen. Mit Ausnahme der Markov-Analyse ist keines der Verfahren in der Lage, den Einfluss unterschiedlicher Wartungsstrategien zu berücksichtigen. Allerdings ist das Markov-Verfahren selbst derart komplex, dass nur eine geringe Komplexität der Aufgabe zu erfassen ist. Weiterhin wird bei diesem Verfahren übersehen, dass Wartungs- und Reparaturmassnahmen einen Einfluss auf das weitere Verhalten des technischen Systems im Hinblick auf die Verfügbarkeit haben.

Nach dem Stand der Technik ist kein Verfahren bekannt, das in der Lage ist, die oben dargestellten hochkomplexen Zusammenhänge so realistisch darzustellen, dass eine zuverlässige Vorhersage der Verfügbarkeit technischer Systeme und damit der notwendigen Redundanz ermöglicht wird.

Darstellung der Erfindung

Ziel der Erfindung ist es also, bei einem Verfahren der eingangs genannten Art die Auswirkungen geplanter und zufälliger Ereignisse auf die Verfügbarkeit eines technischen Systems realitätsnah zu berücksichtigen.

Erreicht wird dies durch die Merkmale des Anspruchs 1.

Kern der Erfindung ist es also, das Unverfügbarkeitsrisiko eines technischen Systems zu ermitteln, und durch eine Anpassung einer Wartungs- und Teileaustauschstrategie zu verbessern, oder aber bereits in der Planungsphase entsprechend Einfluss zu nehmen. Erfindungsgemäss wird hierzu ein Funktionsdiagramm des technischen Systems erstellt, welches das Zusammenspiel der unterschiedlichen Komponenten wiedergibt. Mittels eines Monte-Carlo-Verfahrens wird sodann die Nutzungsdauer des Systems in einer Folge von Simulationsdurchläufen simuliert. Dabei wird aufgrund der vorgegebenen Unverfügbarkeitswahrscheinlichkeit jeder Einzelkomponente bei jedem Simulationsdurchlauf entschieden, ob ein Unverfügbarkeitsereignis eingetreten ist. Die vorgegebenen funktionellen Zusammenhänge des Gesamtsystems in der Simulation ermöglichen, die Auswirkungen des Versagens einer einzelnen Komponente auf die Verfügbarkeit des gesamten Systems zu ermitteln, und die Folgenschwere einer Unverfügbarkeit zu bestimmen. So kann beispielsweise festgestellt werden, ob ein sofortiger Ausfall des Gesamtsystems die Folge der Unverfügbarkeit einer Komponente ist, und wie lange diese Unverfügbarkeit des Gesamtsystems andauern wird. Auf diese Weise können die Unverfügbarkeiten des Systems über alle Simulationsschritte, anders ausgedrückt die Nutzungsdauer des Systems, mit der jeweiligen Folgenschwere gewichtet aufsummiert werden, woraus sich der Verlauf des Unverfügbarkeitsrisikos des Systems über seiner Nutzungsdauer ergibt.

Ein wesentlicher Vorteil des Verfahrens ist also, neben der reinen Unverfügbartkeitswahrscheinlichkeit auch die Folgenschwere einer Unverfügbarkeit – Dauer, Resourceneinsätz zur Behebung eines Schadens, geplante oder ungeplante Unverfügbarkeit – berücksichtigen zu können. Ein weiterer grosser Vorteil des erfindungsgemässen Verfahrens ist, dass eingetretene Ereignisse im folgenden Simulationsschritt berücksichtigt werden.

Kurze Beschreibung der Zeichnung

Fig. 1 zeigt beispielhaft den Verlauf der Unverfügbarkeitswahrscheinlichkeit eines technischen Systems als Funktion der Nutzungsdauer für unterschiedliche Wartungsintervalle.

Fig. 2 zeigt den aus diesen Strategien resultierenden Verlauf des Unverfügbarkeitsrisikos über der Nutzungsdauer.

Fig. 3 zeigt schematisch einen Ablaufplan zur Ausführung der erfindungsgemässen Verfahrens.

Weg zur Ausführung der Erfindung

Die technische Lehre der Erfindung sei vorweg an einem sehr greifbaren Beispiel illustriert. **Fig.** 1 zeigt den zeitlichen Verlauf der Unverfügbarkeitswahrscheinlichkeit eines technischen Systems, wobei ein zur Minimierung des Unverfügbarkeitsrisikos, und somit zur Erzielung eines maximalen Grades der Aufgabenerfüllung, mithin also einer optimalen Resourcenausnutzung, am besten geeignetes Wartungsintervall bestimmt werden soll. Ganz konkret könnten dies die Wartungsintervalle eines Kraftfahrzeuges sein.

Zum Zeitpunkt $t = 0$ wurde soeben eine Wartung beendet. Die Erfahrung lehrt, dass unmittelbar nach Eingriffen in ein technisches System die Unverfügbarkeitswahrscheinlichkeit leicht erhöht ist. So wurde beispielsweise die Ölablassschraube nicht richtig festgezogen sein, oder ein Keilriemen gegen einen aus fehlerhafter Produktion ausgetauscht. Unmittelbar nach einer Wartung sinkt die Unverfügbarkeitswahrscheinlichkeit recht schnell auf ein Minimum, und steigt dann aufgrund fortschreitenden Verschleisses wieder an. Bei der nächsten Wartung ist der Kraftwagen dann mit Sicherheit unverfügbar; die Unverfügbarkeitswahrscheinlichkeit wird also zu 1. Im Beispiel wurde weiterhin berücksichtigt, dass beim längsten Wartungsintervall die Unverfügbarkeitsdauer grösser ist.

Wie ersichtlich ist, wird die kumulierte Unverfügbarkeitswahrscheinlichkeit bei der Wartungsstrategie 1 am kleinsten.

Diese Betrachtung ist jedoch zur Beurteilung der Aufrechterhaltung der Wertigkeit des technischen Systems

"Kraftwagen" absolut unzureichend. Denn die Folgen einer Panne bei Nacht und Regen auf einer einsamen Landstrasse sind ungleich höher zu bewerten als die Folgen eines geplanten Werkstattbesuches: Im letzteren Fall kann zur Aufrechterhaltung der technischen Funktion "Fortbewegung" problemlos im Voraus der Einsatz einer Ersatzresource, beispielsweise Bus oder Fahrrad, zurückgegriffen werden. Weiterhin müssen bei der Beurteilung des Unverfügbarkeitsrisikos des Gesamtsystems auch immer die Unverfügbarkeitswahrscheinlichkeit und Folgenschwere der Unverfügbarkeit einzelner Komponenten berücksichtigt werden: Ein defektes Scheibenwischerblatt mag zwar vergleichsweise wahrscheinlich sein, jedoch werden die meisten Fahrer dies ohne Werkstattaufenthalt beheben können, weshalb die Folgenschwere praktisch Null ist. Der Ausfall der Lichtmaschine bedingt zwar in den meisten Fällen einen Werkstattaufenthalt, die Wahrscheinlichkeit, mit der Batterie die nächste Werkstatt oder gar das Reiseziel zu erreichen, ist jedoch relativ gross. Die Folgenschwere einer ausgefallenen Ölpumpe oder gar eines Bremsenversagens muss hingegen sehr hoch beurteilt werden.

Weiterhin ist bei Schäden, die einen Werkstattaufenthalt bedingen, zu unterscheiden, ob sie unmittelbar die technische Aufgabenerfüllung des Systems verhindern, oder ob die Funktion in Grenzen aufrechterhalten wird, ein notwendiger Werkstattaufenthalt also im Voraus geplant werden kann.

All diese Faktoren sind durch ein Verfahren, zur Ermittlung der Verfügbarkeit eines technischen Systems sinnvoll miteinander zu verknüpfen.

Wird für das angeführte Beispiel das resultierende kumulierte Unverfügbarkeitsrisiko ermittelt, so zeigt **Fig.** 2, dass Wartungsstrategie **2** recht nahe am Optimum liegt.

Die Ermittlung des Unverfügbarkeitsrisikos wird erfindungsgemäss durch eine Monte-Carlo-Simulation vorgenommen. Diese simuliert die Nutzungsdauer eines technischen Systems als Folge von Zeitschritten. Vor jedem Zeitschritt findet eine Zufallsentscheidung statt, ob eine Komponente versagt oder nicht, wobei die Häufigkeit der Versagensereignisse von einer vorgegebenen Versagenswahrscheinlichkeit einer Komponente abhängt. Ein Versagen wird – auch in Abhägigkeit von der Signifikanz der Wirkung einer Komponente im System – mit der Folgenschwere gewichtet. Das heisst, anhand der Funktionsblöcke kann entschieden werden, ob ein Versagensereignis eine unmittelbare Unverfügbarkeit des Gesamtsystems zur Folge hat, und welcher Resourceneinsatz notwendig ist, um den Defekt zu beheben. Dabei ist die Monte-Carlo-Simulation in der Lage, auch die Zusammenhänge sehr komplexer Systeme mit vertretbarem Aufwand zu erfassen. Dabei kann ein nicht behobener Schaden Einfluss auf die folgenden Simulationsschritte nehmen, ebenso wie der Austausch einer Komponente.

Ein Beispiel für den erfindungsgemässen Verfahrensablauf ist im Flussdiagramm in **Fig.** 3 dargestellt. Zunächst müssen der Simulation die Daten für die Unverfügbarkeitswahrscheinlichkeit der einzelnen Komponenten als Funktion der Einsatzdauer einer Komponente zur Verfügung gestellt werden. Danach kann mit der Simulation ab einem Bezugszeitpunkt begonnen werden. Das nachfolgende Funktionsdiagramm muss das Zusammenspiel der einzelnen Komponenten des technischen Gesamtsystems wiedergeben, und besteht allgemein aus Blöcken, die parallel, in Serie oder als k-aus-n-Logiken zusammengesetzt sind. Es ist hierbei nützlich, die einzelnen Komponenten ihrer Funktion entsprechend in hierarchisch gegliederte Untersysteme aufzuteilen. Im Ausführungsbeispiel sind dies zwei Subsysteme A und B, wobei Subsystem B einen unmittelbaren Einfluss auf die Funktion des technischen Gesamtsystems hat, Subsystem A keine zumindest unmittelbare Wirkung.

In einem ersten Schritt wird überprüft, ob die Subsysteme A und B verfügbar sind. Ist Subsystem B unverfügbar, oder ist eine Wartung geplant, tritt ein Unverfügbarkeitsereignis ein, das jeweils mit der Folgenschwere der jeweiligen Unverfügbarkeit gewichtet wird. Im Falle eines Stillstandes können nun mehrere Parameter verändert werden, beispielsweise wird die Einsatzdauer einer ausgetauschten Komponente wieder zu Null gesetzt, oder deren Unverfügbarkeitswahrscheinlichkeit verändert, falls die Komponente gegen eine verbesserte Version ausgetauscht wurde. Schliesslich kann das Unverfügbarkeitsrisiko als Funktion der Zeit dargestellt werden.

Die Versagenswahrscheinlichkeit einer einzelnen Komponente, die bei dem erfindungsgemässen Verfahren ausreicht, um die Verfügbarkeit des Gesamtsystems zu ermitteln, kann beispielsweise aus Erfahrungen im Feldeinsatz empirisch gewonnen worden sein. Bei konstruktiven Veränderungen sind die Auswirkungen der ergriffenen Massnahmen zu extrapolieren. Die Daten der Zuverlässigkeit der Einzelkomponenten sind also mit vergleichsweise hoher Genauigkeit bekannt. Bei einem sinnvollen Aufbau der Funktionsblöcke des erfindungsgemässen Verfahrens kann die Gesamtverfügbarkeit zu einem bestimmten Zeitpunkt der Nutzungsdauer in sehr engen Grenzen spezifiziert werden.

Ein solches Verfahren ist selbstverständlich bereits bei der Auslegung eines technischen Systems äusserst nützlich. So kann beispielsweise die Redundanz verschiedener Subsysteme im Funktionsdiagramm dargestellt und die Auswirkungen einer Veränderung der funktionalen Zusammenhänge bereits in der Design-Phase beobachtet werden. Weiterhin kann auch die Verfügbarkeit eines beliebig komplexen Gesamtsystems in Abhängigkeit der Wartungszyklen und des Austauschs von Verschleissteilen ermittelt werden, um somit eine verbesserte Resourcenausnutzung zu ermöglichen. Werden als Folgenschwere Ereigniskosten eingesetzt, ermöglicht dies wiederum eine Vorhersage von Projektkosten.

Die technische Lehre des Erfindungsgegenstandes besteht zusammenfassend darin, dass dem Fachmann eine Möglichkeit offenbart wird, bereits in der Planungsphase hochkomplexe technische Systeme so auszulegen, dass die Erfüllung der technischen Aufgabe des Systems mit grösstmöglicher Effizienz aufrechterhalten wird. Dabei können auch negative Wirkungen einer steigenden Systemkomplexität und eine daraus resultierende schlechtere Resourcenausnutzung frühzeitig erkannt und vermieden werden. Da zufällige und geplante Ereignisse erfasst und entsprechend ihrer Folgenschwere gewichtet werden, können weiterhin optimale Zeitpunkte für Wartung und Teileaustausch bestimmt werden. Ebenso kann, wenn eine Komponente in einer verbesserten Ausführung zur Verfügung steht, der beste Zeitpunkt für deren Austausch mit Hilfe des erfindungsgemässen Verfahrens ermittelt werden.

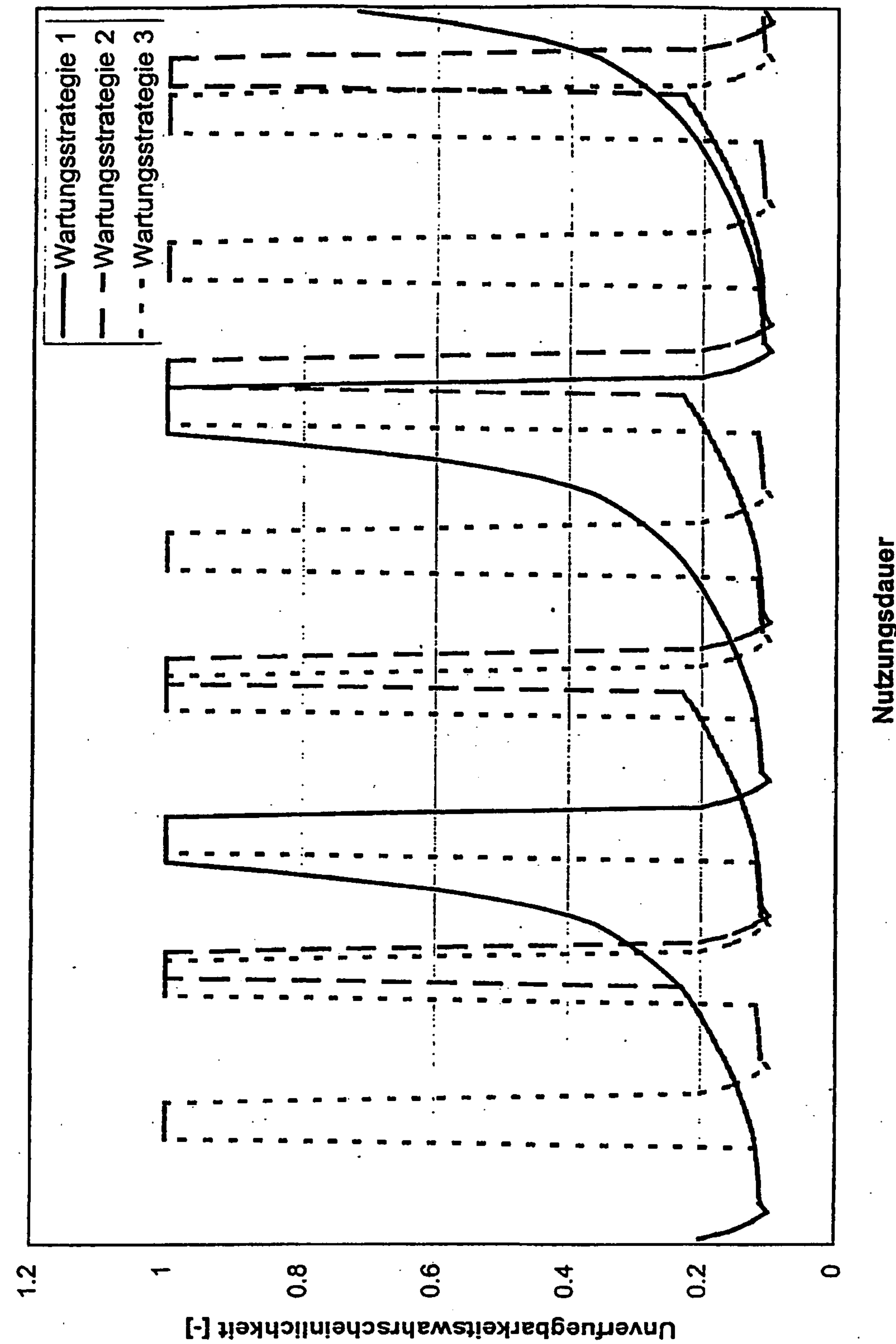
Insgesamt weist das erfindungsgemässe Verfahren somit den Weg zu einer Optimierung der Verfügbarkeit technischer Systeme und einer verbesserten Resourcenausnutzung.

Patentansprüche

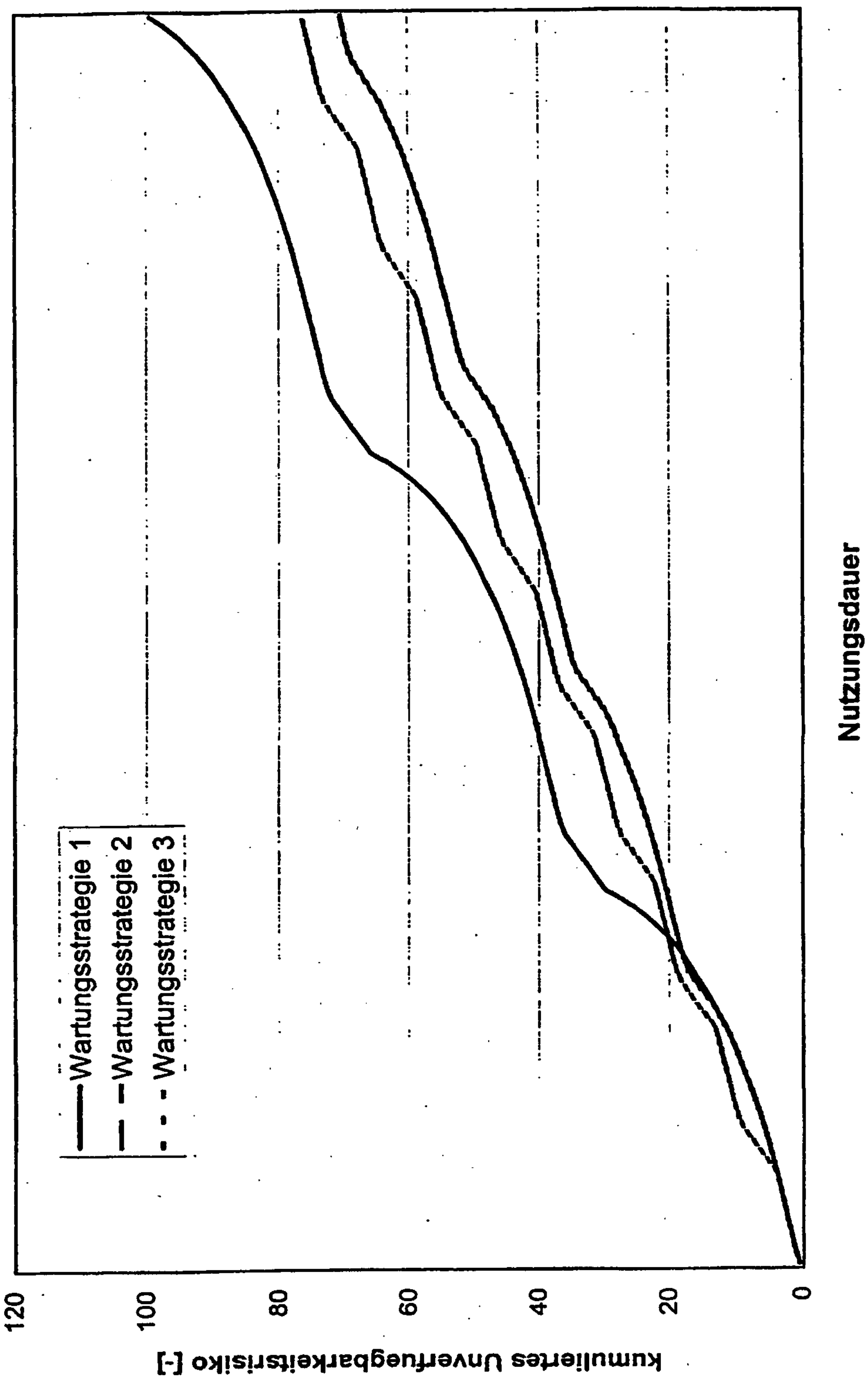
Verfahren zur Ermittlung des Grades der Aufgabenerfüllung technischer Systeme unter Berücksichtigung einer Wartungs- und Teileaustauschstrategie, **dadurch gekennzeichnet**, dass in einer Monte-Carlo-Simulation ein funktionelles Zusammenwirken einzelner

Komponenten des Systems durch Funktionsblöcke dargestellt wird, welche Funktionsblöcke die Auswirkungen des Versagens einer Komponente auf die Verfügbarkeit des Gesamtsystems wiedergeben, dass die Veränderung der Verfügbarkeit des Systems aufgrund zufälliger und geplanter Ereignisse mit Faktoren gewichtet wird, welche Faktoren aus einer Folgenschwere einer Unverfügbarkeit gebildet werden, welche Folgenschwere die Planbarkeit der Unverfügbarkeit und den zur Behebung notwendigen Resourceneinsatz beinhaltet, dass der zeitliche Verlauf der Systemverfügbarkeit in Abhängigkeit von einer Wartungs- und Teileaustauschstrategie ermittelt wird, dass eine teilempirisch ermittelte Lebensdauer unterschiedlicher Systemkomponenten berücksichtigt wird, und dass die Auswirkung von Ereignissen wie Reparatur- und Wartungsmassnahmen auf den weiteren zeitlichen Verlauf der Systemverfügbarkeit berücksichtigt wird.

Hierzu 3 Seite(n) Zeichnungen

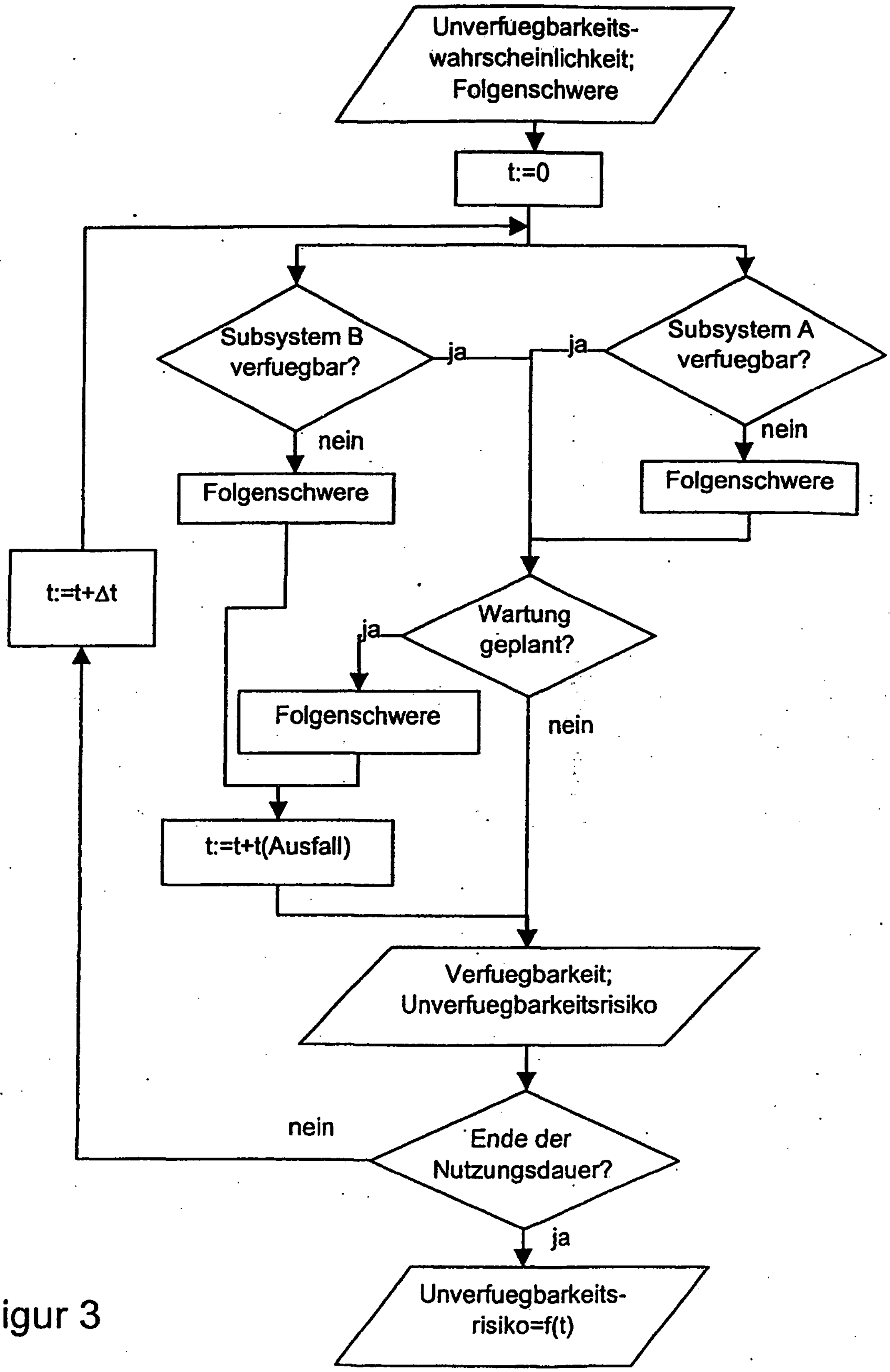

Wartungsstrategie 1
Wartungsstrategie 2
Wartungsstrategie 3
Unverfuegbarkeitswahrscheinlichkeit [-]
1.2
1
0.8
0.6
0.4
0.2
0
Nutzungsdauer

Figur 1


Wartungsstrategie 1
Wartungsstrategie 2
Wartungsstrategie 3
120
100
80
60
40
20
0
kumuliertes Unverfuegbarkeitsrisiko [-]
Nutzungsdauer

Figur 2

Nummer: DE 198 48 094 A1
Int. Cl.[7]: G 06 F 17/60
Offenlegungstag: 20. April 2000



Figur 3

4

Enterprise Modeling and Simulation: Complex Dynamic Behavior of a Simple Model of Manufacturing

Simulating a structurally simple model of a manufacturing enterprise revealed complex dynamic behavior. Enterprise modeling and simulation provided estimates of end-of-life inventory and order delivery performance based on interactions of forecast quality, quoted product availability, material procurement and safety stock policies, vendor lead times, product life cycle, and part commonality. An unexpected result was that end-of-life inventory can exist even under ideal environmental conditions. Prospective applications of these methods include estimating the effects of incremental improvements, verifying impacts of process changes, and generating enterprise behavior information.

by M. Shahid Mujtaba†

Can we understand the potential impacts of process changes? Can we quantify the expected amount of improvements and benefits? Can we anticipate the effects of environmental changes? Can we predict the effects and side-effects of making changes? And can we do all these before taking action and making major resource commitments?

We suggest that the answer is yes to all these questions, and the means is enterprise modeling and simulation.

The purpose of this paper is to show how enterprise modeling and simulation research activities at HP Laboratories can be applied to predict system behavior and gain insights using sound engineering and scientific principles and techniques before implementing the new solution at the level of the business enterprise.

In this paper, we first discuss modeling and simulation technology in broad terms to provide background and context. We then describe one model, the Simple Model, in detail, and present the insights gained from running simulations on that model and analyzing and displaying the results. An unexpected insight was that end-of-life inventory existed at the end of the product life cycle even though the method for computing safety stocks should theoretically have resulted in none when customers ordered exactly according to forecast. Other interesting insights were that high inventory levels can occur when actual orders come in too high or too low with respect to forecasts. In other words, forecast quality has a major impact on some of the metrics under consideration. We then describe the current state of enterprise modeling and simulation, future research directions, and possible application areas, including process reengineering on page 86. In the appendixes we include more detailed explanations and sufficient technical details of the model to permit the results to be duplicated by other researchers. A glossary of terms and a summary of the values for different experiments are provided for quick reference on pages 85 and 95. The evolution of enterprise modeling and simulation activities at HP Laboratories and the place of the Simple Model in those activities provides a historical context and is described on page 90.

Modeling and Simulation

Extensive literature exists on the simulation modeling process, for example Chapter 1 of Law and Kelton,[1] Chapter 1 of Pritsker,[2] Chapter 6 of McHaney,[3] and Law and McComas.[4] The general consensus is that the purposes of the simulation modeling process are to define a problem clearly and to develop a model as a tool to understand and solve that problem.

"Modeling and simulation have become endeavors central to all disciplines of engineering and science. They are used in the analysis of physical systems where they help us gain a better understanding of the functioning of our physical world. They are also important to the design of new engineering systems where they enable us to predict the behavior of a system before it is actually built. Modeling and simulation are the only techniques available that allow us to analyze arbitrarily nonlinear systems accurately and under varying experimental conditions."[5]

"The facility or process of interest is usually called a *system*, and in order to study it scientifically we often have to make a set of assumptions about how it works. These assumptions, which usually take the form of mathematical or logical relationships, constitute a *model* that is used to try to gain some understanding of how the corresponding system behaves."[1]

Thus, a model is a conceptual abstraction of an existing or proposed real system that captures the characteristics of interest of the system. Modeling is the process of building the abstraction (model).

† Author can be reached at email address mujtaba@hpl.hp.com.

"If the relationships that compose the model are simple enough, it may be possible to use mathematical methods (such as algebra, calculus, or probability theory) to obtain exact information on questions of interest; this is called an analytic solution. However, most real-world systems are too complex to allow realistic models to be evaluated analytically, and these models must be studied by means of simulation."[1]

"Simulation is the use of a model to develop conclusions that provide insight on the behavior of any real world elements. Computer simulation uses the same concept but requires that the model be created through programming on a computer."[3]

In general, modeling and simulation are useful when system prototyping is too costly or time-consuming, seriously disruptive, or simply impossible. They are useful for exploring proposed system changes by providing performance estimates of a proposed system or of an existing system under some projected set of operating conditions. A simulation model or set of models can provide an experimental testbed on which to try out new ideas or concepts, since it is cheaper to experiment in the laboratory than on the real system.

Our premise is that these techniques applied to enterprise processes could help predict the behavior of the organization more quantitatively than repeated assertion or the application of mental models.

Enterprise Modeling and Simulation

We define enterprise modeling as the process of building abstractions or models of three primary functional components of an enterprise: manufacturing, marketing, and R&D (research and development) for the purpose of gaining insight into the interactions between these functions and the interaction of the enterprise with other enterprises. The complexity of the enterprise and the large number of people who have ownership of different parts makes it difficult for a single individual to grasp a detailed understanding of all the components. There is a limit to the level of complexity and the means to share and communicate it with others that can be carried in the head of a single individual.

Many process changes and decisions are based on implicit mental models in the heads of decision makers or advocates. Mental models[6] are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action. Very often, we are not consciously aware of our mental models or the effects they have on our behavior.[6] Generally mental models assume that there are a small number of factors in cause and effect relationships. The problem with mental models is the difficulty of communicating them, checking their consistency, and combining the mental models of different people. It is very difficult to estimate the effects of interacting factors and to combine mental models into a larger-scale composite model that incorporates the insights, knowledge, and understanding of many individuals.

One means of merging different viewpoints is the use of Hierarchical Process Modeling,[7,8] which provides an explicit, graphical representation of the process with which individuals can agree or disagree. Experience in applying Hierarchical Process Modeling[9] to the building of enterprise models suggests that the result is a repository for knowledge of the processes we are studying. During its creation, team members develop a common understanding of the dynamics of model behavior through interaction with one another and with the model. The result is an explicit model that reconciles differing points of view and a reusable model that serves as a foundation on which to build future models.

There is an awareness that a model can be used to embody knowledge of a system rather than be used as a tool.[10] For example, Funke[11] states that at the Boeing Company, simulation has provided "a forum for the collection of process operating rules and assumptions in one medium as a basis to develop the model" of a process or system.

Other ongoing works on the application of models to embody knowledge at the enterprise level of manufacturing operations include TOVE[12] and CIM-OSA.[13,14] Pardasani and Chan[15] describe the expansion of an infrastructure for creating simulation models based on the ISO reference model for shop floor production standards to create enterprise models.

In applying the process of enterprise modeling and simulation we need to engage in activities of modeling in the large (with "model as knowledge") where the major issues of interest are communication and documentation, team coordination, modularity and large model development, and multimodel organization, instead of modeling in the small (with "model as tool") where the issues of interest are top-down design, informal and formal program specifications, simplification and elaboration, and validation and verification.[10]

In modeling the manufacturing enterprise, the primary area of focus is the manufacturing function, which includes, in addition to the traditional production and shop floor functions, the production and material planning, material management, and order processing functions. In traditional modeling and simulation applied to the manufacturing domain, computer simulations have been applied to the production floor or machine shop level to study machine utilization and production and material flows and buffers. These methods together with traditional operations research methods have helped reduce inventory on the production floor and cut build times to a level where these are small compared to the other parts of the system. Enterprise modeling and simulation expand the scope so that traditional modeling and simulation are components in the enterprise modeling and simulation system.

Enterprise modeling and simulation indicate the impact of proposed improvement efforts at the enterprise level before the changes are made. The "simulation" in enterprise modeling and simulation is the process of running the model in a computer to understand the behaviors over time under different operating conditions and circumstances. It will help us identify leverage points and indicate where we can expect to get the most impact for a given investment or change.[16]

According to Senge,[6] "The real leverage in most management situations lies in understanding dynamic complexity, not detail complexity." He suggests that most systems analyses focus on detail complexity (that is, a large number of variables), not dynamic complexity ("situations where cause and effect are subtle, and where the effects over time of interventions are not obvious"). We suggest that enterprise modeling and simulation help in understanding dynamic complexity, and in addition provide the framework for slowly expanding the detail complexity.

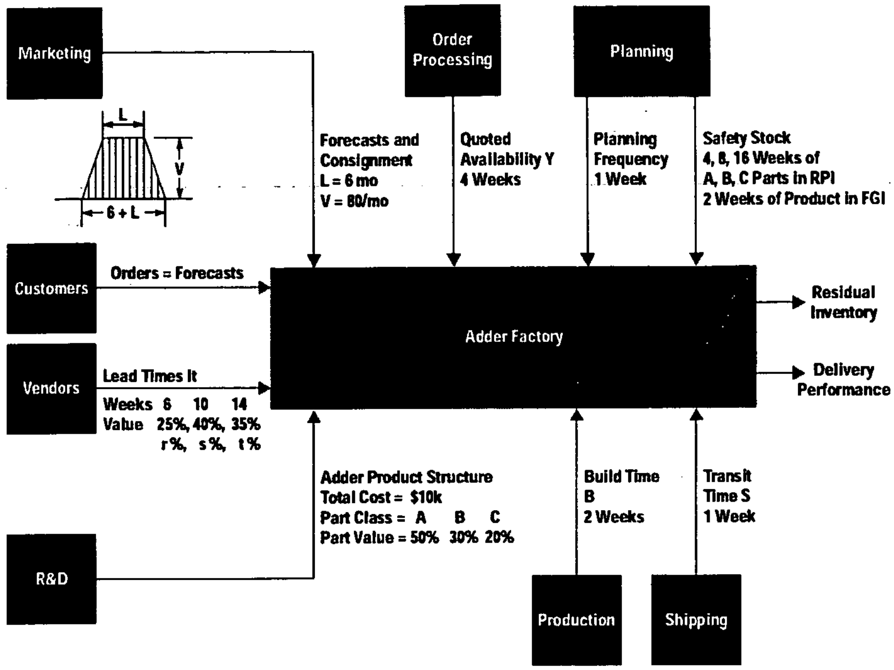


Fig. 1. Diagram of the Simple Model for the nominal case experiment.

Modeling and simulation at the enterprise level are showing increasing levels of activity. For example, a recent article in Fortune magazine[17] discusses business-oriented economics that focuses on what economists call "the firm" and the rest of us call "the company" as the unit of analysis. (Traditional microeconomics, by contrast, is concerned with markets and prices. It looks at the economy or at an industry, but rarely peeks inside the individual enterprise.) Fortune cites the example of Merck's finance team, which built a completed model and subjected it to Monte Carlo simulation analysis.

The Simple Model

The Simple Model (shown with capital letters because of its importance in this paper), was one in a series of models developed at HP Laboratories (see page 90). The Simple Model was named because of its structural simplicity, but as subsequent descriptions will show, it exhibits dynamic behavior that is complex and not intuitively obvious until it is explained. Expressed in terms used by Senge,[6] the Simple Model is a tool for understanding dynamic complexity using a model with very low detail complexity.

The Simple Model was commissioned to abstract a real manufacturing facility with greatly simplified assumptions, such as a single product with a one-level bill of materials and a trapezoidal order demand pattern. The purpose of the model was to explore the relationship between different factors and metrics used in manufacturing. Although the model can generate data on many different metrics, this paper will focus on two main metrics: (1) inventory levels and write-off at the end of the product life cycle and (2) customer satisfaction metrics. We will first describe the structure and assumptions of the Simple Model and then show the results of running the model under different conditions.

Conceptual Description

Fig. 1 shows conceptually the Simple Model of a factory producing a product called Adder.† Marketing specifies a trapezoidal order forecast profile for customer orders, and the number of consignment units (defined as demonstration units used in the sales offices). R&D specifies the Adder product structure. Order processing quotes a product availability of four weeks. Production determines that the build time is two weeks, and shipping states that transit time for sending the product to the customer is one week. We assume that the production and shipping processes are under sufficient control that they do not vary from these constant numbers.

The problem assumes that the values of class A, B, and C parts in the Adder product make up 50, 30, and 20 percent, respectively, of the product material cost. In valuing the finished product, labor cost is small enough to be factored into the material cost, and the value of the product is the sum of values of its parts. In addition, we assume that the values of 6-week, 10-week, and 14-week lead time parts make up 25, 40, and 35 percent, respectively, of the product cost, that the vendors deliver the parts exactly on time, and that there are no rejects because of defective parts.

These characteristics are reflected in Table I, which shows the value of each part category. There are a large number of unit costs and part quantity combinations that satisfy the above constraints. The actual bill of materials used for the model is shown in Table II.

The length of the longest lead time among the parts is 14 weeks for parts A.3, B.3, and C.3. Allowing a build time of

† There was a little bit of whimsy in naming the product. The author selected the name from a fairy tale in which somebody ordered the biggest adder available, expecting it to be an adding machine. When the box was opened, out popped a snake. Snakes, of course, was an internal HP code name for a class of workstations.

Table I
Simple Model Adder Product Structure

(a) Product Structure by Part Value

Part	Value	Part	Value	Part	Value
A.1	$1250	B.1	$750	C.1	$500
A.2	$2000	B.2	$1200	C.2	$800
A.3	$1750	B.3	$1050	C.3	$700

(b) Part Value by Part Class Safety Stock

Class	Parts in Class	Value	% Value	Safety Stock
A	A.1,A.2,A.3	$5000	50%	4 weeks
B	B.1,B.2,B.3	$3000	30%	8 weeks
C	C.1,C.2,C.3	$2000	20%	16 weeks
	Total	$10,000	100%	

(c) Part Value by Lead Time

Lead Time	Parts	Value	% Value
6 weeks	A.1,B.1,C.1	$2500	25%
10 weeks	A.2,B.2,C.2	$4000	40%
14 weeks	A.3,B.3,C.3	$3500	35%
	Total	$10,000	100%

Table II
Adder Bill of Materials

Part	Quantity	Unit Cost	Value in Product
A.1	1	$1250	$1250
A.2	1	$2000	$2000
A.3	1	$1750	$1750
B.1	1	$750	$750
B.2	1	$1200	$1200
B.3	1	$1050	$1050
C.1	1	$500	$500
C.2	1	$800	$800
C.3	1	$700	$700

two weeks and transit time of one week means that the period from the time parts A.3, B.3, and C.3 are ordered in the manufacturing enterprise to the time that the product using those parts is received by the customer is 17 weeks. This means that the policy of waiting for customer orders before we order parts from our vendors will lead to an order-to-delivery time of at best 17 weeks.

To quote availability of four weeks requires us to order material and plan production before we receive customer orders. The best information we have on current and past customer behavior is actual orders, and the best information we have on future customer orders is the order forecast.

Given that we want quoted availability to be less than the sum of material delivery, production, and product delivery times, we need to plan ahead of time how much to build based on order forecasts. This decision on how much to build in future weeks is the responsibility of production planning, which each week computes the number of units to be started in future weeks.

Forecasts of future customer orders are estimates; customers may order more or less than forecasted. In the event that customers order less, we should have no problem meeting the demand if we build to meet the forecast. However, if customers order more, we might run out of product. To allow for this contingency production planning must specify that we need to build a few more units and carry them in a stock of finished goods inventory (FGI). The amount of extra product to be carried is the safety stock, and depends on many factors including the average expected order level, the expected fluctuations in orders, and how much we want to allow for contingencies. A high safety stock level will protect us from low forecasts, but requires a greater investment in inventory. One way of specifying inventory levels is to use a measure related to number of weeks of forecasted demand. In the case of this model, we assume that production planning specifies two weeks of 13-week leading average forecast as target FGI safety stock.

The discussions for FGI safety stock are also applicable for raw material. There must be enough raw material on hand when the time comes to build the product. To allow for excess demand from the production line because of high customer demand, and for late deliveries by vendors, we need to order some extra material. This extra amount is determined by material planning and is the target raw parts inventory (RPI) safety stock. The amount of RPI safety stock can be determined in different ways. One way is to use part classification.

In practice, part classification indicates the relative importance of a part and hence the attention it receives. Since class A parts are reviewed more frequently, a smaller quantity is carried than for the B or C parts. In our model, part class determines the amount of material safety stock to be carried in weeks, and all parts are reviewed weekly by material planning. For A, B, and C parts the target RPI safety stock is 4, 8, and 16 weeks, respectively, of the 13-week leading average forecast. The 13-week leading average forecast and the FGI and RPI target safety stocks are discussed in greater detail under "Target Safety Stock," below.

Fig. 2 shows the trapezoidal product order forecast supplied by marketing. The demand during each week of a four-week month is constant. The demand builds up over three months,

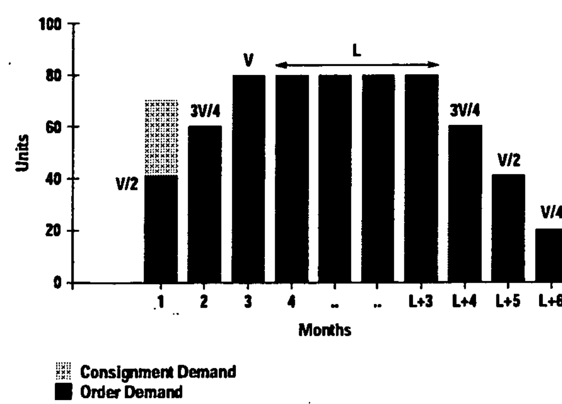


Fig. 2. Adder order forecast and consignment demands in units.

remains constant for L months, and then reduces to zero over three months, so the total product life is L+6 months. In the first month, some units are required for consignment purposes. The mature monthly demand V is 80 units, and the total amount of inventory for consignment is set at 1.5 weeks of projected mature demand, or 30 units. In our experiments we used a baseline value of 6 months for L. This order forecast results in a lifetime total of 780 units, or a total forecasted production cost flowthrough (PCFT, see Glossary, page 85) of $7.8 million, exclusive of the 30 consignment units.

Of the many performance metrics for the system during the product life cycle, the three main ones of interest are the end-of-life inventory, which needs to be disposed of or written off, the shipment and delivery performance, and the inventory during the product life cycle.

Detailed Description

The fundamental description of the Simple Model of the enterprise and the primary flows and dynamic components that interact with it over time are shown in Fig. 3.

Entities External to the Enterprise. Customers send orders to the manufacturing enterprise. In the simulation each order for a single unit is sent individually to the manufacturing enterprise. The orders go into the backlog of the manufacturing enterprise, and at some point a shipment fulfilling each order is delivered to the customer. Customers have the expectation that the time between ordering and receipt of delivery is the quoted availability, but are willing to wait indefinitely for orders.

The manufacturing enterprise sends orders for each part to the respective vendor, shown collectively in Fig. 3 as vendors. The shipment of physical parts arrives at some time in the future determined by the lead time for the part. Ideally the time between the issuance of an order and receipt of the material (parts) should be the lead time quoted by the vendor, and for all the runs in this paper, this will be the case.

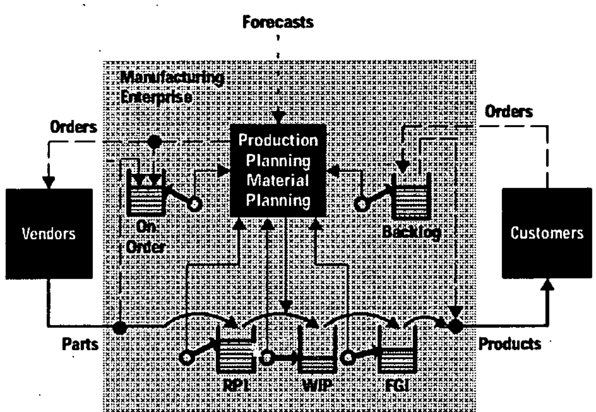


Fig. 3. Material, order, and information flows of the Simple Model simulation. The heavy solid lines represent the flow of physical material, the long-dash lines represent the flow of information related to individual orders, and the short-dash line represents the flow of periodic order forecasts. The containers represent the accumulation of physical material or orders, the pointers represent levels of the quantities in the containers, and the light solid lines from the containers represent this status information being transmitted to the planning function. The light solid line from the planning function represents a control signal flow that regulates the amount of material flowing from RPI to WIP and ultimately to FGI.

Functions Internal to the Enterprise but External to Manufacturing. Periodically, *marketing* provides forecasts of customer orders in future periods. Each forecast is a list of the quantity of products that are estimated to be ordered in subsequent periods. In practice, forecasts are updated periodically and estimates for the same month in the future can vary from month to month. In the model, the forecast is used to compute the shipment plan, and to compute the 13-week leading average forecast for computing FGI and RPI safety stocks. *R&D* (not shown in Fig. 3) provides a bill of materials (BOM) that defines the product structure. Since the BOM does not change during the simulation, we do not show the R&D function.

Processes Internal to the Manufacturing Function. This section should be read in conjunction with Figs. 1, 2, and 3.

Order processing accepts orders and keeps track of all outstanding orders received from customers, and keeps a running total of the quantity of products required in the backlog. In addition, it prioritizes the orders by the ranking criterion, which in this model happens to be first-in, first-out (FIFO), into a ship list. The backlog level is provided to the production planning function. The prioritized list of orders and the quantity that needs to be shipped in the current period are provided to shipping.

Shipping fills and ships the orders on the ship list that order processing provides. From the point of view of the manufacturing enterprise, the duration between receipt of customer order and delivery of the shipment at the customer site should be the time period specified as the quoted availability. Filling an order is attempted no earlier than necessary to satisfy the quoted availability taking transit time into account. An order is filled and shipped only if at the time of the attempt the number of units in FGI is greater than zero. In other words, shipping's objective is to fill outstanding orders that need to be filled and not to try to maintain FGI at some level. This means that the actual order-to-delivery time for a particular order will depend on whether units are available to fill the order at the time the order is due to be shipped. If units are not available, the order will have a higher priority for being filled in the next period because of the FIFO rule used to establish the ship list.

Production planning computes the current shipment plan and build plan. It computes the current shipment plan from the current order forecasts and current order backlog to attempt to satisfy the quoted availability. It then computes the current build plan from the shipment plan, build time, current FGI, current WIP, and FGI safety stock.

To come up with a suitable build plan, production planning must know about the characteristics of the system it is trying to control, that is, it must have a model of the system that it uses for doing its computation. An important aspect of the computation is to take into account the number of units already in process rather than relying only on the number of units of product required. Such a model is generally a mathematical or analytical model, and the formulation is described in Appendix I. The build plan for the current period is used

Glossary of Terms and Abbreviations

Abbreviations

A/F. Actual-to-forecast ratio. This is the ratio of the actual orders received to the forecasted orders. Normally expressed as a percentage. A/F greater than 100% implies that actual orders came in higher than forecasts, that is, forecasts were low or demand was high. A/F less than 100% implies that actual orders came in lower than forecasts, that is, forecasts were high or demand was low.

BOM. Bill of materials. A description of the components that go into an assembly and their respective quantities.

- Single-level BOM. The components are raw materials fabricated or manufactured elsewhere (i.e., purchased parts).
- Multiple-level BOM. The components are other assemblies and purchased parts.

EOL. End of life (end of product life cycle).

FGI. Finished goods inventory.

RPI. Raw parts inventory. Raw material in stores waiting to be processed.

WIP. Work in process. Raw material on the production line being assembled into the final product.

PCFT. Production cost flowthrough. Dollar value of production passing through the manufacturing enterprise. Because of the assumptions underlying the Simple Model, in this paper PCFT is synonymous with shipments from the manufacturing enterprise.

Terms

Backlog. Products ordered by customers but not yet shipped.

Build Time. The time required for completion of the product when all the parts are available.

Committed Inventory. The total inventory to which the manufacturing enterprise is currently committed. It is the sum of the on-order material and the on-hand inventory.

Consignment Inventory. Inventory in the sales offices and for demonstration purposes.

End-of-Life Inventory. The amount of inventory left over at the end of the product life cycle, that is, when no more orders are backlogged or outstanding for the product. EOL inventory includes leftover unused RPI, unshipped units in FGI, and consignment inventory. In general, material and products left over at the end of the product life cycle are not useful for anything else and must be written off.

Forecast Quality. Qualitative description of the amount of deviation of actual customer orders from forecasted orders. The ratio A/F described above is one way to quantify forecast quality. Forecast quality is best for A/F = 100%, and gets worse as A/F moves away from 100%.

Lead Time. The time between placement of an order to the vendors and receipt of the material.

On-Hand Inventory. All physical inventory that is owned by the enterprise. It is the sum of RPI, WIP, and FGI.

On-Order Inventory. Same as on-order material.

On-Order Material. The total amount of material for which orders are currently open and which will eventually be received from vendors. It increases each time a new order is issued and sent to the vendor, and decreases each time a shipment of parts is received from the vendor.

On-Time Delivery. Measures whether the order is delivered to the customer within the quoted availability. When described in units or dollars, it represents the units or dollar value of the deliveries that are delivered within the quoted delivery time. When described as a percentage it represents the percentage of on-time deliveries with respect to the total deliveries.

On-Time Shipments. Products that were shipped to customers within the quoted availability minus the transit time, that is, those shipped to arrive in time to satisfy the quoted availability.

Order Backlog. The total amount of outstanding orders from customers that have not yet been shipped. It increases each time a new order is received from customers, and decreases each time an order is shipped to customers.

Order-to-Delivery Time. The time period from order issue to order delivery at the customer site.

Order-to-Ship Time. Time period from order receipt to order shipment at the manufacturing enterprise.

Orders Delivered. Orders that have been delivered to customers.

Orders Shipped. Orders that have been shipped to customers.

Product Life Cycle. The general shape of the increase, leveling off, and decrease in order volume for the product. We assume here it is trapezoidal.

S and S-Plus. S is a language and interactive programming environment for data analysis and graphics developed at AT&T Bell Laboratories. S-Plus is a product version of S that is sold and supported by Statistical Sciences, Inc.

to trigger the start of the appropriate number of units in the current period.

Material planning uses the BOM to generate a material consumption plan for each part that can support the build plan. It then uses the material consumption plan, on-order material, RPI, RPI safety stock, and part lead times to determine the material ordering plan, that is, how much of each part to order in the current and future weeks. Details of the computation of the consumption and ordering plans are given in Appendix I.

Material ordering sends orders for the appropriate amount of each part in the current week to the vendors. As each order is sent, the on-order material for that part increases.

Raw material stores (not shown in the figures) receives and stores incoming material in RPI and provides material to production when requested. As it receives deliveries from vendors, it sends information about the shipment to on-order material which is reduced by the amount of the shipment received.

Production receives a build plan and requests as much material as required from raw material stores to build the number of units required. Only complete sets of parts are drawn from stores, that is, if one or more parts are not available in sufficient quantities, all parts are drawn partially. For example, if the build plan calls for 10 units to be built, and there are only 5 units of part A.3 and more than 10 units each of the other parts in RPI, only 5 units of of each part will be drawn and sent to WIP, and only 5 units can be started. The objective of raw material stores is to fill requests for material if possible, and not to maintain RPI at any particular level. The mathematical derivation of the number of units actually started subject to the available material is given in Appendix I.

Enterprise Modeling and Simulation Applications in Reengineering

Process reengineering as defined by Hammer and Champy in their book, *Reengineering the Corporation*,[1] is "the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service, and speed." It is being applied at an increasing rate by three kinds of companies: those in deep trouble, those not yet in trouble but whose management has the foresight to see trouble coming, and those in peak condition with no discernible difficulties whose management is ambitious and aggressive. These three categories cover a large number of companies. The impact is on processes with throughputs measured in the billions of dollars.

Reengineering is pervasive, controversial, and disruptive, and has different interpretations. CSC Index, whose chairman is Champy,[1] states that even though they pioneered the practice of reengineering, they are startled by how widespread the phenomenon has become. Their survey results[2] based on 497 large companies in the U.S.A. and another 124 in Europe show that 69% of the U.S. companies and 75% of European companies are already reengineering (average completed or active initiatives in excess of 3). More than half of the rest were planning to launch an initiative over the next 12 months or were discussing one.

Hammer and Champy[3] mention three kinds of techniques that reengineering teams can use to help them get ideas flowing: boldly apply one or more principles of reengineering, search out and destroy assumptions, and go looking for opportunities for the creative application of technology.

A sampling of the literature reveals that redesign is influenced by the past experience of the reengineering team and the recommendations of reengineering consultants. Ultimately, many redesign decisions are made on speculation based on implicit mental models, convincing arguments by vocal proponents for change, sheer optimism, blind faith, or desperation.

A major concern is the uncertainty of predicting outcomes. Radical redesign and new ideas bring the possibility of boundless gain or tremendous loss. While assumptions are being searched out and destroyed ruthlessly, it should not be forgotten that some assumptions are rooted in scientific principles which cannot be ignored with impunity no matter how highly enthusiastic or motivated the reengineering team.

Enterprise Modeling and Simulation

Some areas suggested by Hammer and Champy[4] for reengineering the corporation include product development from concept to prototype, sales from prospect to order, order fulfillment from order to payment, and service from inquiry to resolution.

The Simple Model described in the accompanying article is a start towards addressing order fulfillment. Modeling and simulating the other processes on the list require different kinds of knowledge acquisition. For example, product development requires more knowledge about the R&D function, sales requires more knowledge about the marketing function, and service has not been considered in the current model, where the focus is on manufacturing.

The following paragraphs describe areas where enterprise modeling and simulation and the enterprise modeling and simulation system may provide value in the reengineering effort.

Identifying Processes

Hammer and Champy[5] suggest that once processes are identified and mapped, deciding which ones require reengineering and the order in which they should be addressed is not a trivial part of the reengineering effort. Typically there are three criteria for making the selection: dysfunction, importance, and feasibility.

Enterprise modeling and simulation provide one way of gaining insight in these areas by generating performance metrics with and without the change under different circumstances. For example, the Simple Model showed the importance of different controllable and uncontrollable factors to the different system performance metrics such as EOL inventory and order-to-delivery cycle times.

After selecting a process for reengineering, an understanding of the current process is crucial. It is necessary to know what the existing process does, how well (or poorly) it performs, and the critical issues governing its performance from a high-level view. This understanding is the prerequisite to redesign. The key is understanding the process rather than completely analyzing it in agonizing detail.

Enterprise modeling and simulation offer at least two ways of obtaining this understanding and possibly showing the cause of the dysfunction. First, the very act of building a consensus model that different people can agree with sheds light on what might not be working. Second, simulating the model will confirm or reject the validity of what is suspected. For example, after building the Simple Model, it was possible to test it in a large number of possible operating conditions to provide understanding of the cause and effect relationships. The first major insight from simulating the model was that what appeared to be a reasonable way of computing safety stock that would go to zero as demand went down actually gave rise to end-of-life inventory even though the demand was forecasted accurately. Enterprise modeling and simulation provide a way of gauging the relative impact of different process changes as a step towards selecting the appropriate subprocess to reengineer, and of quantifying the amount of prospective improvement.

Enterprise modeling and simulation can show the prospective impact of infeasible changes. In simulating the proposed reengineering changes, even if they are infeasible, the results will indicate if there is any promise in further consideration of a particular direction. For example, it is clearly not feasible to have zero build

The required material is drawn from RPI and goes into WIP where it remains for the duration of the build period. After that, the completed units go into FGI.

Target Safety Stock. Inventory is the amount of physical material, and ideally the enterprise would like to maintain it at or close to zero in RPI and FGI, and only carry it in WIP when raw material is being converted into final product. In practice, to reduce the effects on production of late vendor deliveries and customer orders coming in higher than forecasts, safety stock needs to kept. In the Simple Model, where vendor delivery time uncertainty is not an issue, to allow for the contingency that customer orders may come in higher than forecast, production planning targets the FGI safety stock to be two weeks of 13-week leading average forecast, and material planning targets RPI safety stock for each part to be the quantity of that part required for the production of the number of weeks specified in Table I(b) of the 13-week leading average forecast.

The 13-week leading average forecast at the end of a particular week in the future is the sum of the order forecasts over the 13 weeks immediately following the particular week divided by 13. This average anticipates trends 13 weeks (one calendar quarter) into the future, increasing as order forecasts increase, and decreasing as order forecasts decrease. In particular, the 13-week average forecast is zero at the end of the product life cycle, which means that any target safety stock expressed in weeks of 13-week leading average will aim for a zero target safety stock level at the end of the life cycle.

Having specified target safety stock in preparation for demands higher than forecasted, what is the impact if customers order exactly according to forecast? The expectation is that actual FGI should be equal to targeted FGI safety stock level, and actual RPI for each part should be equal to targeted RPI safety stock level for that part.

time for products and zero transportation times for shipments in the real world, but setting those values to zero in the model indicates the theoretical maximum benefits of these actions, and the magnitude of the results provides a data point for decisions on how much investment to put on driving these two times to zero instead of on other opportunities.

Furthermore, by showing the time behavior of the changes, enterprise modeling and simulation can show when actions can be expected to take effect. Inertia is a property of most systems, reflected in the time taken to respond to external influences or changes. Most physical systems are predictable in this respect, but the time behavior for organizational systems such as the enterprise is less predictable simply because it is not understood as well. Enterprise modeling and simulation help to increase the predictability of system behavior given that we know something about the system's structure and the behavior of its components. While immediate improvement for reengineering is the desired goal, enterprise modeling and simulation can show the length and causes of delays in obtaining the desired result.

Exposing and Challenging Assumptions

Hammer and Champy suggest that we question assumptions.[6] Enterprise modeling and simulation require assumptions to be stated explicitly during the model building process to reconcile differences in points of views. Challenges and disagreements on the validity are with respect to clearly stated assumptions rather than differences in opinions resulting from differences in mental models of different individuals. For example, the production planning and material procurement processes used in the Simple Model are expressed mathematically in Appendix I. If these are accepted as rational methods of planning, then there is no question or debate on the values of the outputs for a given set of inputs. If processes expressed mathematically are not acceptable as rational methods of planning and an alternative method is proposed, then that alternative method can certainly be tried, and the results compared with the previous method. The debate and challenge for improvement becomes one of improving the logic of planning rather than one revolving around the meaning of words and labels or one on how the model should behave based on past experience or speculation.

The approach advocated by Hammer and Champy suggests that changes be made by understanding the problem and devising the solution. This is central to modeling and simulation in addressing problems in the realm of the enterprise. Enterprise modeling and simulation offer a way of testing and verifying that given the current knowledge, the results of the simulation do not exhibit any obvious flaws before the process is implemented.

Role of Technology

Hammer and Champy devote a whole chapter to discussing the essential enabling role of information technology, and assert that modern state-of-the-art information technology is part of any reengineering effort. They caution that the misuse of technology can block reengineering altogether by reinforcing old ways of thinking and old behavior patterns, and that equating technology with automation does not result in reengineering.

We suggest that the application of enterprise modeling and simulation is a creative application of a well-understood technology to the processes of the enterprise. The technology of modeling and simulation has been applied to fields such as product design and the design of physical systems, but is only now beginning to be applied creatively in analyzing the processes of the enterprise. What enables the creative application of modeling and simulation is the tremendous increase in computational power. In this respect, we would like to suggest another rule along the lines of the rules described in reference 1.

Old Rule: Decisions regarding process changes are based on mental models and analysis of historical data.
Disruptive Technology: Enterprise modeling and simulation.
New Rule: Decisions regarding process changes are based both on historical data and analysis of computer simulated behavior of explicit models with explicit assumptions that show the prospective consequences of different actions under a large number of operating circumstances.

Conclusion

Reengineering is a philosophy of renewal and rapid, discontinuous, and drastic change in the way corporate enterprises do their work, which brings with it uncertainty and fear of the unknown future. It is disruptive and controversial, and there is as yet no agreement that successes outnumber failures. During the implementation, "People focus on the pain of the present and the joy of the past. They forget about the pain of the past and the joy of the present."[7] However, given that it is occurring on such a wide scale, we suggest that application of enterprise modeling and simulation can increase the chances for success by (1) quantifying the potential benefits of the reengineered process in an explicit, defensible way, (2) illustrating the transition between the pain of the present and the joy of the future, and (3) showing the possible outcomes of current actions, thereby making the future more predictable and less surprising to those most affected by it.

References

1. M. Hammer and J. Champy, *Reengineering the Corporation: A Manifesto for Business Revolution*, Harper-Collins Publishers, Inc., 1993.
2. *State of Re-Engineering Report, Executive Summary*, CSC Index, 1994.
3. M. Hammer and J. Champy, *op cit*, p. 146.
4. M. Hammer and J. Champy, *op cit*, p. 118.
5. M. Hammer and J. Champy, *op cit*, p. 122.
6. M. Hammer and J. Champy, *op cit*, p. 145.
7. J. Kornbluth, "The Prophet of Pain," *Worth*, Vol. 3, no. 6, July/August 1994, pp. 80-81.

Simple Model Simulation

The Simple Model described above represents a simple process design for a manufacturing facility that is subject to simulation. On the surface, the design appears to be reasonable and adequate, and in fact is based on representative data and characteristics of the process. However, simulations will show some unexpected behavior, as well as the envelope of the possible behaviors.

The Simple Model was executed on an evolving system called the EMS system, which consists of two parts: the simulation engine part and the data analysis and display part. The simulation engine has continued to develop with each model that we have studied. It captures and abstracts processes in the enterprise. The simulation engine is an object-oriented, enhanced discrete event simulation software system.

The initial implementation of the simulation engine part of the EMS system was the Manufacturing Enterprise Simulator on the TI Explorer II.[9] The current implementation runs on HP 9000 Series 700 workstations at the Manufacturing Systems Technology Department of HP Laboratories. The implementation language is the Common Lisp Object System (CLOS).[18] The simulation engine has been implemented in CLOS provided by three different vendors: Franz, Inc.,[19] Lucid, Inc.,[20] and Harlequin, Ltd.[21] Models subsequent to the Simple Model (see page 90) were large enough to stress the limits of all three implementations. Graphical output was produced using S-Plus. Further details of the history and development of the EMS system are given in reference 9. The initial version of the Simple Model was implemented within a week based on the full order-to-ship model[22] (see page 90). It then took successive refinement and a tremendous amount of time to analyze the results.

For the reader familiar with discrete event simulation, details of the similarities and differences in concept between this implementation and conventional discrete event simulation are discussed in reference 9. In general, orders and shipments

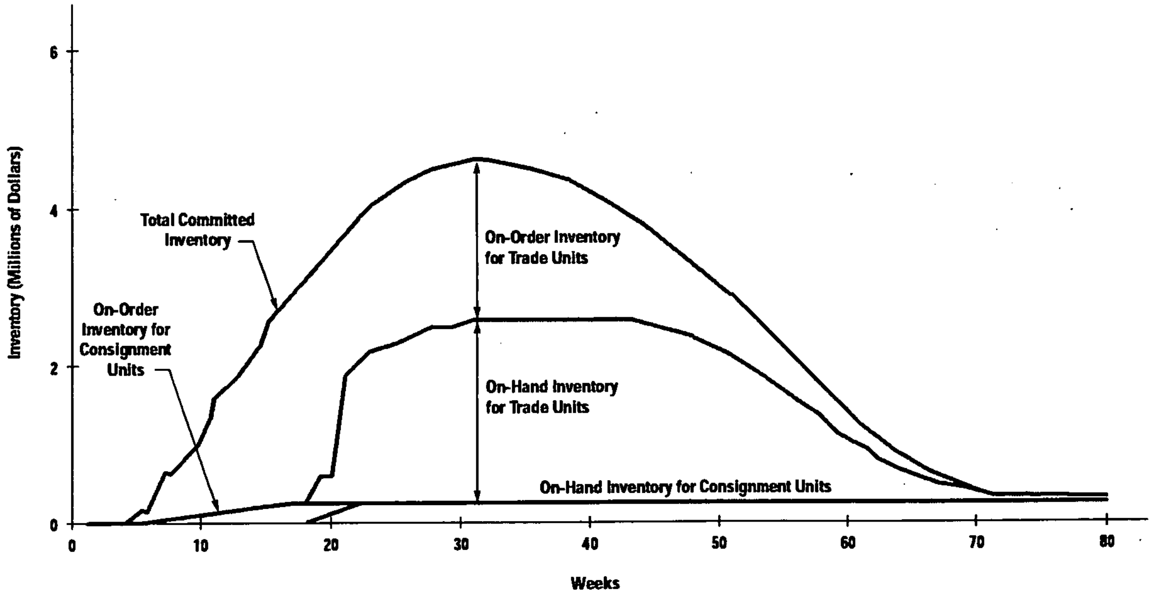


Fig. 4. Nominal case inventory components as functions of time. The experimental conditions are shown in Fig. 1.

are modeled as the entities of discrete event simulation. Backlog, on-order material, RPI, WIP, and FGI are modeled as queues. Customers and vendors are modeled as source-sink combinations of orders and material and vice versa. Production is modeled as an activity.

The production and material planning functions, which are essentially information processing and decision making functions, are implemented as mathematical models embedded in the simulation. The information generated by these planning functions determines when and how many units of product to start building and how many units of material to order. Thus, we can think of the Simple Model as an analytical mathematical model embedded in a discrete event simulation model. The analytical model (formulation given in Appendix I) dictates how the simulation model should behave in the same way as the planning functions dictate how operations should be handled in reality. The simulation model is the reflection of physical reality and reflects the behavior of the physical system that is told what to do.

There are two aspects of uncertainty: bias and variance. Most simulation models focus on variance and assume bias (offset) to be zero. While the EMS system supports the ability to simulate the model under stochastic conditions, in the runs described in this paper, variance is always zero and the emphasis of the analysis is on the situation in which bias can be nonzero.

Each run represents one combination of inputs and parameters of the system, and the traditional statistical analysis of means and confidence levels is not directly applicable for the analysis of these runs. While process variances are important considerations in a system, the motivation of this work was to identify the first-order effects of the various factors, considering the variances as second-order effects.

Details of the timing of the event sequence are shown in Appendix II.

Experimental Results

Experiment 0: The Nominal Case

The nominal case experiment assumes ideal conditions for testing the model. The purpose is to establish model baseline behavior and offer face validation by verifying that results are consistent with intuition and the observed behavior of the real system. Initial conditions for committed inventory and backlog are set to zero. A warmup period of five months (20 weeks) allows material to be ordered and received before customer orders arrive on week 21. The last customer orders arrive on week 68. Order forecasts are consistent with the trapezoidal profile already defined, and while they are generated weekly, they do not change from week to week. Week 21 corresponds to the first week of month 1, and week 68 corresponds to the last week of month L+6 in Fig. 2. Production begins during week 19 to ensure units in FGI at the end of week 21. The computation of FGI and RPI safety stock levels is assumed to apply only for weeks after week 20. Up to and including week 20, the required safety stock level is set to 0.

Time Response of On-Hand Inventory and On-Order Material. Fig. 4 shows inventory levels measured in dollar terms over time. The two bottom regions show the on-order material and on-hand inventory for consignment units. There is a gradual buildup of on-order material, which is rapidly transformed into on-hand inventory over four weeks, followed by a flattening out (since the consignment units are never shipped). The middle region shows on-hand inventory for trade or shippable units, which is the sum of RPI, WIP, and FGI. The upper region shows the on-order material commitment for trade units. The top surface of the graph shows how total material commitment changes over time.

Inventory Investment. Committed inventory at the end of week 20, before the first customer order arrives, is approximately

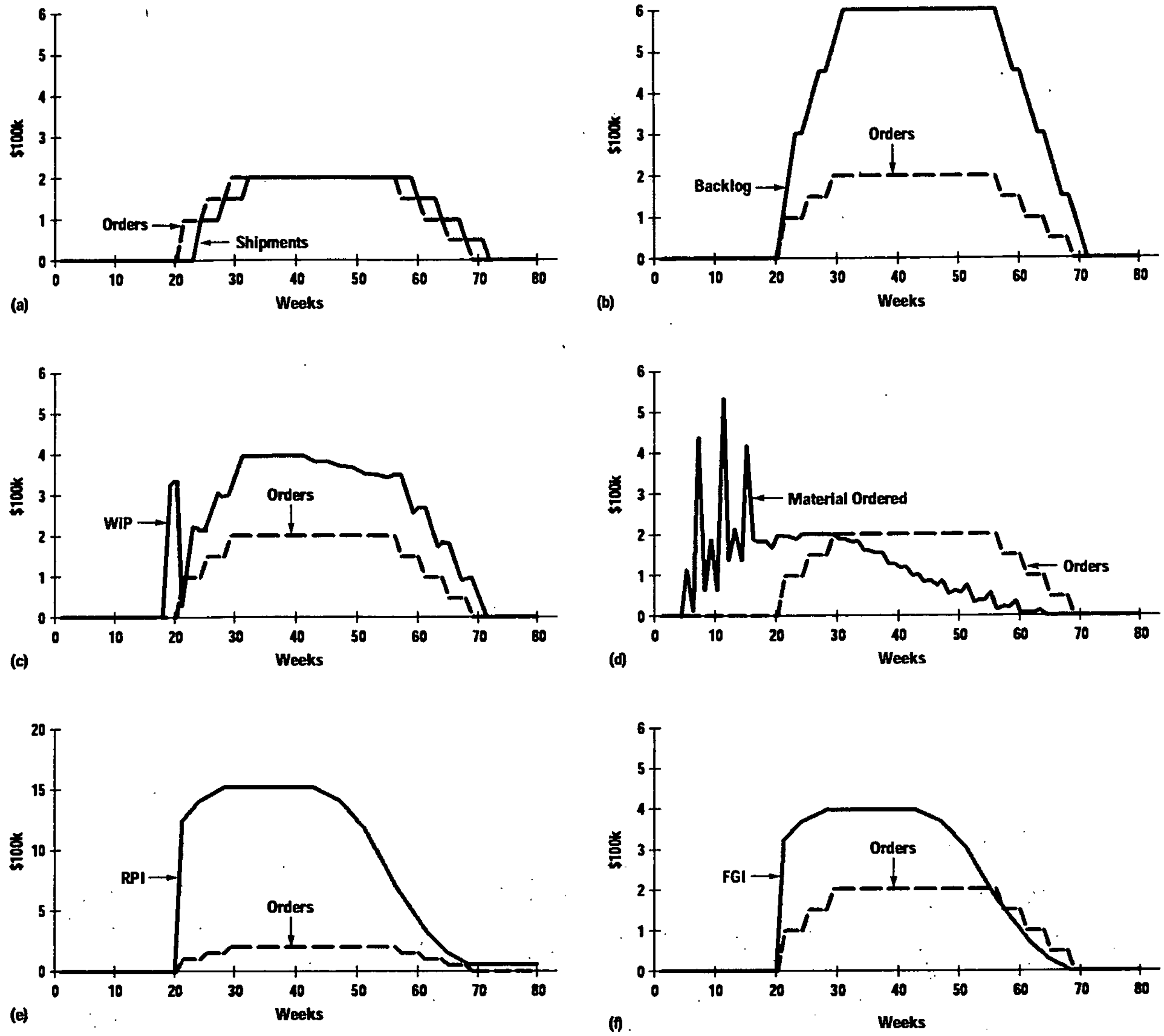


Fig. 5. Metrics as functions of time for the nominal case. (a) Shipments and orders. (b) Backlog and orders. (c) WIP and orders. (d) Material ordered and orders. (e) RPI and orders. (f) FGI and orders.

$3.5 million. If orders to vendors cannot be cancelled, this $3.5 million commitment must be disposed of if we decide to cancel the product before the first customer order arrives.

During the mature part of the life cycle of the product, the on-hand inventory is approximately $2.5 million and the total committed inventory is approximately $4.7 million. To support shipment levels of $200,000 a week requires $4.7 million of committed inventory (23.5 weeks of steady-state PCFT) and $2.32 million of on-hand inventory (11.6 weeks of steady-state PCFT), both of which include $300,000 of consignment units (1.5 weeks of steady-state PCFT). Details of the computations verifying these numbers in the simulation are given in Appendix IV-2. The maximum inventory exposure over the life cycle is $4.7 million.

The EOL consignment inventory of $300,000 reflects the amount of potential write-off because we did not dispose of the consignment units. The EOL nonconsignment inventory for trade units is reflected in the tail of the graph, and its value is approximately $64,000. If the material cannot be consumed any other way, there is an EOL write-off of $64,000 of nonconsignment inventory and $300,000 of consignment inventory for a PCFT of $7.8 million under ideal conditions of perfect forecast quality and on-time vendor delivery.

Time Response of Other Metrics. Fig. 5 shows other time series metrics in comparison to orders received. The shipment profile (Fig. 5a) is identical to the order profile but shifted in time by three weeks. This is because the four-week availability and one-week transit time require three weeks of order-to-ship time for on-time delivery.

Steady-state backlog (Fig. 5b) is $600,000, or three weeks of orders. Again, this is because the four-week availability and one-week transit time result in orders staying in backlog for three weeks, that is, the current backlog is the sum of the last three weeks of orders.

Enterprise Modeling and Simulation Research at HP Laboratories

Our work at HP Laboratories on enterprise modeling and simulation is an outgrowth of the factory modeling project, which began in early 1987. While we were working in the area of robotic automation for manufacturing, we began to appreciate the complexity of the geographically distributed, multientity marketing, manufacturing, and distribution operations necessary for HP to remain competitive. We also realized that there were very few tools available to help understand, design, and operate these complex systems.

Having been involved in product design with the evolving use of CAD and CAE tools, we thought that there was an opportunity of potentially tremendous magnitude for applying similar technologies to the design and operation of the factory and business systems used to market, manufacture, and distribute products. In an effort to capitalize on this opportunity, we began identifying the primary elements of a single factory and building our preliminary order-to-ship model that spanned all major activity from the receipt of an order to its shipment.

Preliminary Order-to-Ship Model

This early model was a vehicle to show the feasibility of applying simulation at a scope larger than a production line, where simulation was beginning to be applied. Developed and proposed for discussion purposes, it was a model to analyze why the order-to-ship time for some products stretched to weeks when the application of modern manufacturing techniques had reduced the build time to a matter of hours. More details on the reasons behind this work are given in references 1 and 2.

Full Order-to-Ship Model

By late 1988 the preliminary model was ready for testing in a real-world context. Data and operational information were provided by a real manufacturing division to help enhance our early model. This process helped to validate the preliminary order-to-ship model and led to the development of the full order-to-ship model.[3] The primary factors considered were order forecast quality, production capacity constraints, supplier lead times, and order filling policies. The primary metrics of interest were order lateness, backlog, and inventory. The model included three distribution centers, one manufacturing entity, and a centralized sales and order entry system. It was configured for one-level bills of materials (BOM), multiline orders, and long life cycle products.

The results of the analysis done with the full order-to-ship model were encouraging; they showed things that were consistent with real-world experiences (e.g., high forecasts led to high inventory and low backlog). The results also provided a view of greater potential by helping to identify areas for future improvement (e.g., the dominant cause of product shortages is long lead time parts coupled with poor forecasts rather than the build time).

While the results of this model were modest, the building and running of this model enabled us to explore some important technologies (i.e., Hierarchical Process Modeling for knowledge acquisition, a discrete event simulation language, SLAM II,[4] and a knowledge-based environment, Knowledge Craft, for system representation and building simulations). Our efforts led to generalized enterprise-level modeling elements and an object-oriented simulator. We also identified some new obstacles (e.g., managing large amounts of simulation data, extracting information) to be overcome in attaining our goals. More details are given in reference 1.

For about a year, no further model development was done, but rather, much effort was put into consolidating what we had learned about the modeling and simulation issues. This effort led to the complete overhaul of our modeling and simulation code while migrating it to the Common Lisp Object System on HP workstations. The power and speed of our system took a quantum leap forward.

Simple Model

With our improved system ready, we were presented with another real-world opportunity to apply our techniques. The Simple Model was proposed as a means of pulling together the main activities, processes and circumstances involved in a manufacturing enterprise. The primary purpose was to understand end-of-life (EOL) inventory and order delivery performance issues. The combined impacts of several environmental factors and operational policies were considered in the

---

Fig. 5c shows an initial spike in WIP preceding the start of orders by two weeks. This happens because the number of units started during week 19 is not only what is to be shipped two weeks later, but also the quantity that must be in FGI (approximately two weeks of orders) at the end of week 21. The WIP levels taper off downwards starting in week 44 towards the latter part of the life cycle because as the desired FGI safety stock level decreases, less production is required than is shipped because some units shipped from FGI do not have to be replenished.

Fig. 5d shows material orders. The three large spikes in material orders are caused by different lead times for parts to fill the targeted RPI safety stock at the beginning of the cycle. Each of the three small spikes corresponds to the different lead time parts for the initial WIP spike. Once the initial spikes are past, the material ordering volume is approximately the same height as the customer orders, except that it is shifted earlier in time, showing that once the system has reached mature demand, material inflow in the form of material ordered is balanced by the material outflow in the form of shipments. Material ordering starts ramping down beginning in week 28 just as the orders reach the maximum demand for this particular set of circumstances.

Fig. 5e shows RPI as a function of time. Notice that the vertical scale is different from the other graphs. The RPI level is 7.6 weeks of PCFT during the mature demand period and starts ramping down in week 44. Fig. 5f shows FGI as a function of time. The FGI safety stock during the mature demand week is two weeks of PCFT, which is the same as two weeks of steady-state orders. The FGI level starts ramping down in week 44.

Inventory Results. The results establish the baseline behavior of a system designed to take contingencies into account when those contingencies do not occur. Appendix IV provides further details for computing some of these results on a theoretical or common sense basis. Some interesting observations can be made. First, EOL inventory and write-off exist even though customers ordered exactly according to forecast and we expect safety stock to go to zero. Second, the level of inventory required to support this level of business can be quantified. Third, long lead time parts make up a greater percentage of the value of parts on order than their percentage in the product structure.

EOL inventory is important for short life cycle products because the inventory cannot be used for anything else and must be written off. In this case it is a result of the way of computing safety stock. It occurs if in the early part of the life cycle too much material is ordered because of high targeted FGI and RPI. For short life cycle products it can be a significant percentage of PCFT. EOL inventory is less of an issue for long life cycle products because the leftover inventory is generally a smaller percentage of total PCFT and excess inventory in early periods can be used at a later time.

analysis. The model, leveraging our earlier work, dealt with a one-level BOM, one factory, one product, and subsequently a family of successive products with common parts and overlapping life cycles.

Our analysis provided some interesting insights, such as certain material procurement and safety stock policies result in EOL inventory even for perfect order forecasts, and with low forecasts, increasing material lead times and planning frequency result in increased EOL inventory. More important, we began to realize that we were onto something that could really have a positive impact for HP. In fact, the business results led to the development of the planning calendar model with the Simple Model as its foundation. We also continued our technical enhancements by connecting the output to S-Plus[5,6] for data analysis and the creation of a Lotus® interface to display output.

Planning Calendar Model

The purpose of the planning calendar model[7,8,9] was to determine the effects of planning cycle times on inventory levels. It required extension of the Simple Model to include production planning and material planning cycle times. It approximated a two-level BOM and multiple assembly sites using a one-level BOM at one site. It used historical forecasts and orders. The primary factors were forecast quality, the length of the planning cycle, and the maximum lead times for parts. The primary metrics of interest were average inventory, delivery performance, and inventory levels at the start of production. The primary technical development was the application of S-Plus data analysis capabilities to the data.

With this model, material lead times had a dominant effect on inventory levels and committed inventory. Historically, forecasts were generally low, so for the historical data given, the planning cycle time used for the particular product had insignificant impact compared to material lead times. There was greater potential for reducing inventory by reducing lead times than by reducing planning time. Low forecasts increased backlogs.

Current Modeling Activities

We are currently finishing an analysis of a single-site manufacturing system where we were looking at how to improve the supplier response time. The challenges in this application include managing a multilevel bill-of-materials and understanding the consequences of long, variable test cycle times. We are also working with sector-level reengineering teams to help understand the consequences of proposed changes and explore alternatives.

Our enterprise modeling and simulation capabilities have evolved considerably from our preliminary order-to-ship model. However, there are still many more interesting challenges to address before we reach our goal of a computer-aided business process design and operation system.

Robert Ritter
Project Manager
Enterprise Modeling and Simulation Project
HP Laboratories

References

1. M.S. Mujtaba, "Simulation Modelling of a Manufacturing Enterprise with Complex Material, Information, and Control Flows," *International Journal of Computer Integrated Manufacturing*, Vol. 7, no. 1, 1994, pp. 29-46.
2. M.S. Mujtaba, "Systems with Complex Material and Information Flows," *Proceedings of the International Conference on Object-Oriented Manufacturing Systems (ICOOMS)*, pp. 188-193.
3. M.S. Mujtaba, *Formulation of the Order-to-Ship Process Simulation Model*, HP Laboratories Technical Report #HPL-92-135, December 1992.
4. A.A.B. Pritsker, *Introduction to Simulation and SLAM II, Third Edition*, Systems Publishing Corp., 1986.
5. *S-Plus Programmer's, User's, and Reference Manuals*, Statistical Sciences Inc., 1992.
6. A.A. Becker, J.M. Chambers, and A.R. Wilks, *The New S Language*, Wadsworth & Brooks/Cole Advanced Books & Software, 1988.
7. C.M. Kozierok, *Analysis of Inventory and Customer Service Performance Using a Simple Manufacturing Model*, Master of Science Thesis for Leaders for Manufacturing (LFM) Program, Massachusetts Institute of Technology, May 1993.
8. K. Oliver, *Simple Model Report*, distributed by email on January 12, 1993.
9. M.S. Mujtaba and R. Ritter, *Enterprise Modeling System: Inventory Exposure and Delivery Performance*, HP Laboratories Technical Report #HPL-94-89, October 1994.

Lotus is a U.S. registered trademark of Lotus Development Corporation.

This nonzero EOL inventory is significant because our safety stock policy targets zero safety stock levels in FGI and RPI at the end of the life cycle. Having observed this phenomenon in the simulation, we were able to show mathematically why the EOL inventory is not zero. The formal derivation of this result is outside the scope of the current paper, but more detailed analysis of the data showed that it is the Class C parts that are left over. The Class C parts will be zero in the case when orders come in as forecasted for the conditions of experiment 0 only if the target RPI safety stock for Class C parts is less than or equal to the 13-week leading average forecast. Also, for the conditions of experiment 0, any part with target safety stock greater than 13 weeks of 13-week leading average forecast will end up with EOL material. The behavior of the amount of Class C EOL material as the number of weeks of target safety stock goes down is given in Appendix IV-3, and an informal explanation showing the reasoning behind the EOL material is given in Appendix IV-4.

The nonzero EOL is a function of the number of weeks of 13-week leading average forecast. Other techniques of computing safety stock, for example using a cumulative leading forecast rather than the 13-week leading average forecast, might lead to different results.

Smoothing WIP and Production. The initial spike in WIP shows how the policy of starting production in week 19 (and not before) gives rise to a spike in capacity demand at the beginning of the product cycle. It could be eliminated by incorporating production capacity constraints into production and material planning or by allowing FGI to build up before the first order comes in (i.e., before week 21). Both of these require production to start before week 19.

Experiment Set 1a: Single Uncontrollable Factor Variation

In the nominal case, the customer order pattern was accurately forecast. We now consider the situation where the actual orders are different from the forecasts.

We assume that customers order according to a constant order forecast profile multiplied by some constant factor Actual/Forecast or A/F. A/F is the ratio of actual orders to forecast orders; its definition is shown in Fig. 6a. In practice, marketing would change the forecasts periodically. Since we were not modeling the forecasting process, we chose the simplifying assumption that although a new forecast is generated every week, it is identical to the forecast generated the previous week.† Here is an example of bias in the order forecast with no variance. The model interpretation is that although estimates were wrong in the past, we expect that future orders will be equal to the original forecast. This is

† This is not a limitation of the model. A user-specified forecast can be accepted by the model. Later models have incorporated historical forecasts. The reason for this assumption was to get a better understanding of the effect of forecast bias. Fluctuating forecast deviations make interpretation harder.

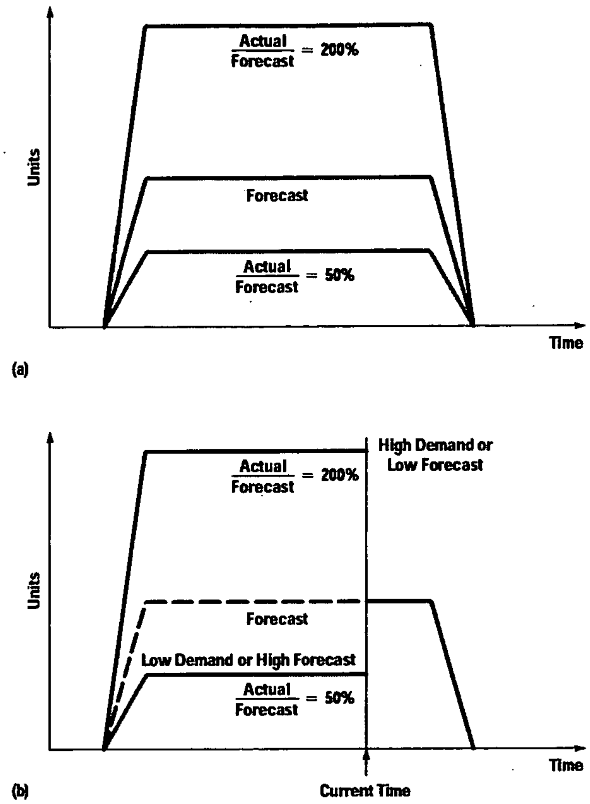


Fig. 6. Definition of A/F. (a) A/F ratio. (b) Actuals and forecasts at the current time.

reflected in Fig. 6b. Actuals came in as shown in the part of the graph to the left of the current time, while the part to the right of the current time line shows the current expectation of future orders.

Clearly, we would expect an effect when A/F is not 100%. If A/F is less than 100%, that is, if forecasts are high, FGI will start to build up, since production planning has directed a larger number of units to be built than are subsequently demanded. Production planning and material planning take this into account and plan to build less and order less material in the future, but the overall material level is higher than when A/F is equal to 100%. On the other hand, if A/F is greater than 100%, that is, forecasts are low, FGI will start to be eaten away because production planning has directed a smaller number of units to be built than are subsequently demanded. Subsequently, production planning and material planning take this into account and raise the production, but since they are always estimating low future demand, we would expect the inventory level in general to be lower than in the case where A/F is 100%. Surprisingly, this intuitive result does not hold, as will be seen later.

We ran the simulations with A/F ranging from 50% to 200% at equal intervals of 25%. In addition, we ran it at smaller intervals in the region of 95% to 125%.

EOL Write-off. A consequence of keeping forecasts identical for all runs is that the consignment profile does not change with respect to A/F. Fig. 7 shows EOL metrics as A/F ranges from 50% to 200%. Note that the changes in value are not constant across the horizontal axis. Fig. 7a shows that total EOL inventory increases as A/F decreases. Fig. 7b shows that the percentage impact is even worse, simply because the write-off is a higher percentage when PCFT, which is directly influenced by A/F, is lower. For low forecasts, that is, A/F greater than 100%, the EOL inventory decreases. For high forecasts, that is, A/F less than 100%, the EOL inventory increases. The lower the A/F, the higher the EOL inventory.

Fig. 7 leads to the obvious conclusion that inventory write-off can be reduced by the strategy of underforecasting orders. However, this is only one side of the story. The complete story is shown in Fig. 8.

Impacts on Time Series of A/F Changes. Fig. 8 shows the impact of A/F changes on different time series measures. *To avoid clutter we will not show inventory for consignment in subsequent time series.* FGI, WIP, RPI, on-order material, and on-hand inventory will refer to the material associated with trade units unless otherwise specified.

All of the graphs in each row of Fig. 8 exhibit identical behavior before week 21. This is to be expected, since before the first orders come in on week 21, the situation is the same for all cases. Only as different amounts of orders come in on or after week 21 is the situation different for different values of A/F.

Fig. 8a shows the order forecasts and actual orders for reference. The ratio of the values of the two lines at any time in the graph is equal to A/F.

Fig. 8b shows the backlog and actual orders time series on the same scale. Notice how the backlog increases spectacularly as A/F goes beyond 125%. Fig. 8c, which displays backlog in terms of mature demand, shows that for an A/F value of

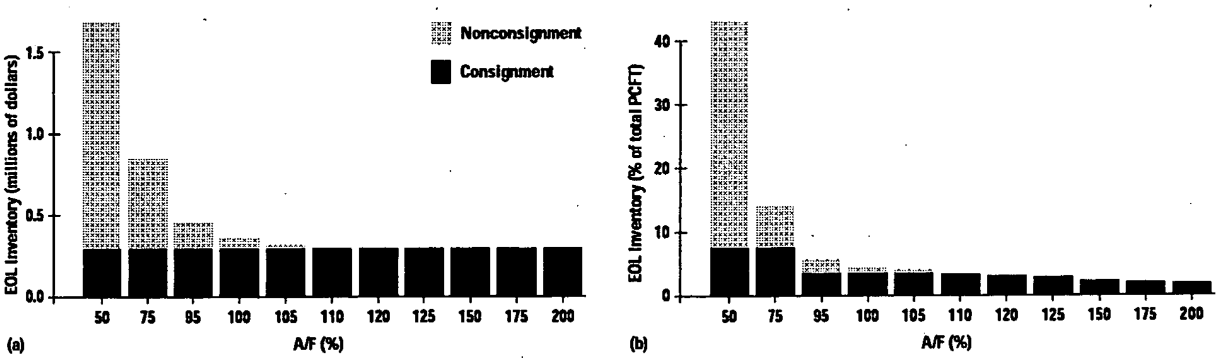


Fig. 7. EOL (end-of-life) inventory for experiment set 1a.

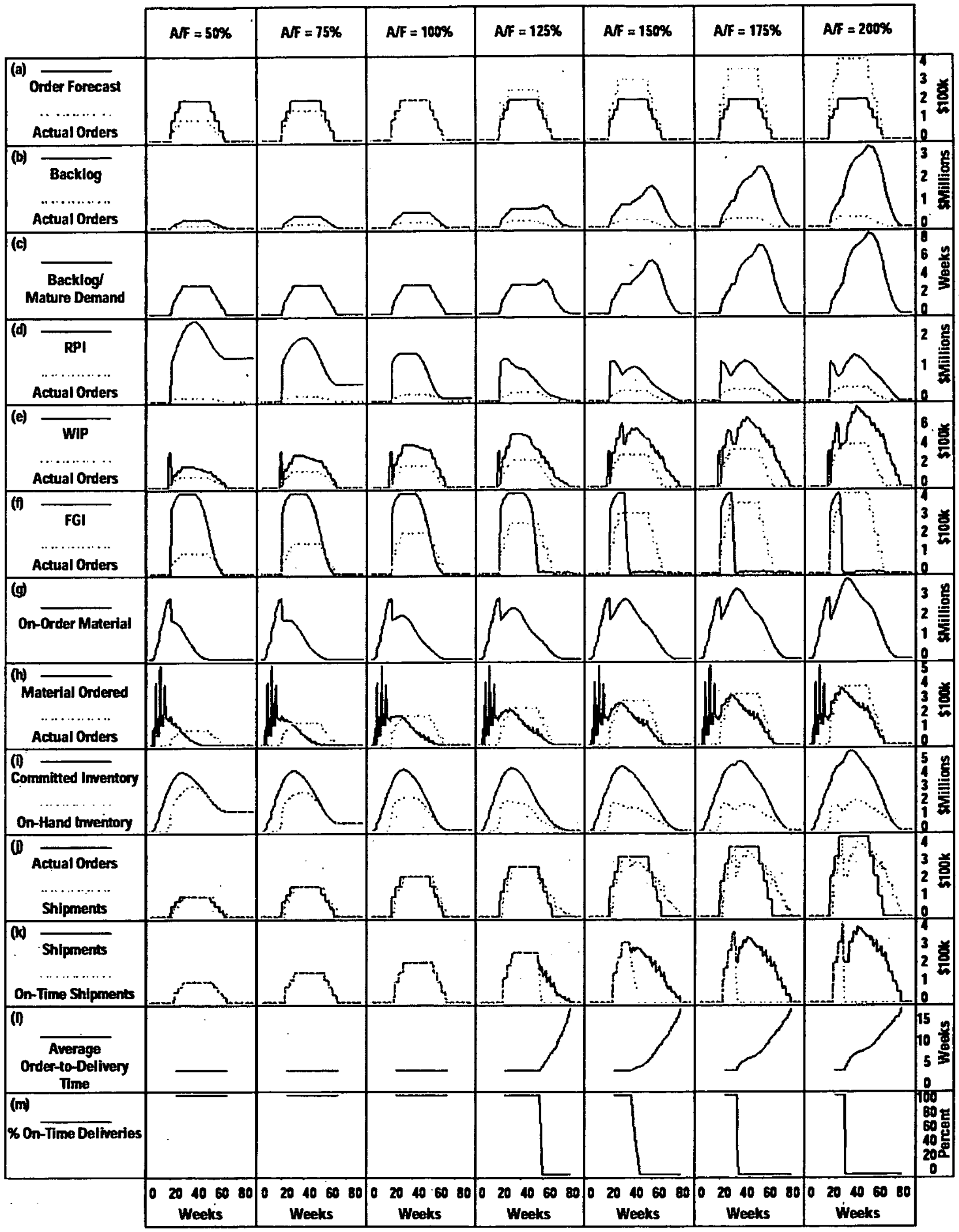


Fig. 8. Time series data for various values of A/F for experiment set 1a.

200% the backlog can be as much as eight weeks of mature demand. Backlog measured in terms of weekly mature demand is constant for low A/F. It increases for high A/F because products cannot be shipped as fast as orders come in.

Fig. 8d shows that the EOL RPI level falls as A/F increases. In addition, the general level of RPI as a function of time falls as A/F increases until A/F is greater than 150%, when the RPI level actually appears to rise as A/F increases. The reason is that because of shortages we order more of all material to build the shortfall in units. The short lead time parts show up first, but cannot be used because of a shortage of the long lead time parts with minimal safety stock. An analysis of the results shows that the critical part is A.3.

Fig. 8e shows that the WIP profile increases as A/F increases. This is expected, since WIP is directly related to the shipments flowing through the system, and the shipments are directly related to orders, which are directly related to A/F.

Remember that this is true only when the production capacity constraint is not reached. If production capacity is only a little greater than forecast, high demands would result in the level of WIP being capped at some limit but spread out over time.

Fig. 8f shows that the FGI level is identical for all values of A/F less than 100%. For A/F greater than 100%, the FGI gets eaten away slowly because the rate of replenishment of new units does not keep up with the shipments because of underforecasting. However, since FGI safety stock levels are based on two weeks of 13-week average forecast and the forecasts used are identical in all the experimental runs, the peak FGI tends to be the same.

Fig. 8g, on-order material, shows initial large spikes for material for RPI and FGI safety stock, followed by a drop after the material for safety stock has been delivered. Subsequently the profile shows an increasing level over time as A/F increases.

Fig. 8h, material ordered, shows the same spikes before week 21 that we have seen before. Again the material ordered versus time increases as A/F increases.

Fig. 8i shows that, in general, committed inventory after week 21 is higher for higher A/F and stretches out farther over time. For lower A/F the committed inventory is lower in the early part of the life cycle, but there is an increase in EOL inventory.

Fig. 8j shows that for A/F less than 100%, shipments follow the order stream nicely. High A/F (high demand) values cause the initial orders to be filled as specified, but subsequently shipments drop off and then catch up. The product shipment over time is smooth when A/F is less than or equal to 100%. When A/F is greater than 100%, during the early part of the life cycle the orders are filled as they come in. As the FGI safety stock is consumed, the shipments fall to the forecasted levels, and then subsequently tend to rise to the actual order levels.

The on-time shipment graphs in Fig. 8k show that initial orders are delivered on time in all cases. For A/F less than 100% (forecasts are high), all orders are delivered on time. For A/F greater than 100% (forecasts are low), initial orders are delivered on time, but subsequent orders are late. As A/F increases beyond 100%, both the percentage and the total dollar value of on-time shipments (and consequently deliveries), go down, and the late orders never catch up. On-time delivery graphs, which are not shown, would be identical to on-time shipment graphs shifted by one week.

As expected, because of the policy of shipping as late as possible, Fig. 8l shows that average order-to-delivery time never goes below four weeks, but increases with time up to 18 weeks as A/F increases to 200%. Fig. 8m, showing the percentage of on-time deliveries, is consistent with Figs. 8k and 8l in terms of on-time deliveries.

How Late Are Late Orders? How late are the late orders and how many orders are delivered on time (namely, within four weeks of being ordered)? These questions are answered in Fig. 9, which shows the dollar volume of deliveries and the order-to-delivery time. For A/F less than or equal to 100% (forecasts high or demand low), all orders are delivered on time. For A/F = 105%, most orders are delivered on time. For A/F = 150% and 200% (forecasts low), some orders are delivered on time, and a large fraction of orders are delivered late. Furthermore, for high A/F values, even though the total volume of shipments is higher, the amount of on-time shipments and deliveries actually goes down. Some orders are delivered as much as 14 weeks late, that is, 18 weeks after receipt of order. This 14 weeks is the upper limit of lateness for this particular model and data configuration. No matter how high A/F gets, orders will never be later than 14 weeks. The explanation for this is given in Appendix IV-5.

Interpretation of Results. In this model, forecasts were not updated on the basis of orders. In reality, when orders are very much under or over forecasts, there will be pressure to change the forecasts. If further information on the forecasting process is available, this can be incorporated into the model. Another study that could be done is to see what happens if we treat the initial orders as early indicators of the whole life cycle, that is, after some period of time, we revise the forecasts so that they more closely represent the volume of actual orders. On the other hand, if the life cycle is very short, it may turn out that revising the forecasts when the first orders come in may not have an impact on system response. We have established a nominal trapezoidal product life cycle, but this could be changed in various ways. It could be stretched out horizontally to increase the life cycle (as is done in subsequent experiments), or vertically, to show a higher level of product demand.

Customers need to receive the products within a reasonably short time, or they might cancel the order. For the model, we assume that customers are willing to wait patiently as long as it takes for the manufacturing facility to produce and ship the products, and that they will not cancel the order.

The purpose of this detailed discussion is to show how changing the one factor, A/F, can have different impacts on different metrics, and how this might affect different parties interested in the outcomes. A/F is partly under the control of customers, and partly under the control of marketing, assuming that greater effort will provide a better estimate of orders. It shows that if A/F is low, order processing and shipping would have excellent performance metrics in getting products out in a timely fashion, whereas material procurement would be in the situation of trying to explain why there is so much material in the plant, and marketing and the plant manager may have to explain why orders are below target. On the other hand, if A/F is high, customers

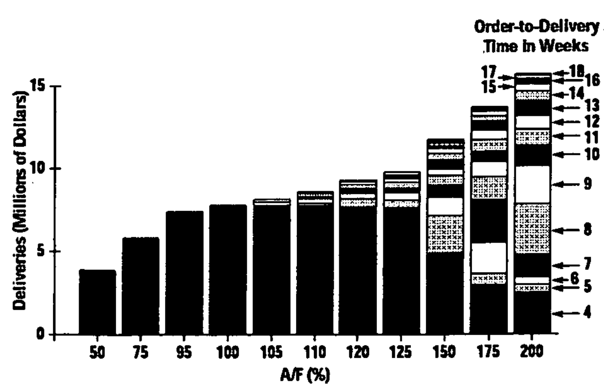


Fig. 9. Deliveries by order-to-delivery time in weeks for experiment 1a.

Table III
Range of Values of Factors for Different Experiment Sets

Factor Description	Parameter Name	Number of Different Values	Values
Actual/Forecast, %	A/F	11	A/F (%) = 50,75,95,*100*,105,110,120,125, 150,175,200
Part Safety Stocks Class A 4K weeks Class B 8K weeks Class C 16K weeks	K	5	K = 0.5,0.75,*1.0*,1.5,2.0
Life Cycle, L+6 months	L	5	L = 0,3,*6*,12,18
Availability, weeks	Y	5	Y = 1,2,*4*,8,12
Percentage value of 6, 10, and 14-week lead time parts in the product (r%,s%,t%)	lt	4	rrr = (100,0,0),rst = *(25,40,35)*, sss = (0,100,0), ttt = (0,0,100)

Experiment 0 (nominal case): Values shown in italics.

Experiment Set 1a (uncontrollable factor A/F varied): Values of A/F varied as shown. Values of other factors same as experiment 0.

Experiment Set 1b (A/F = 100%, controllable factors varied): Values of factors other than A/F varied as shown in turn. Values of other factors same as in Experiment 0.

Experiment Set 2 (dual-factor experiments): Values of factor A/F and one other factor varied in turn. Values of other factors same as in Experiment 0.

Experiment Set F (all factors varied): Values of all five factors varied as shown.

will be screaming for products, order processing and shipping will be trying to placate angry customers, production will be under pressure to put out products faster, and material procurement will have to explain the perpetual shortage of raw material A.3 while other material is piling up.

Experiment Set 1b: Controllable Factors Varied with 100% A/F

We next look at the effect of changing the factors over which the manufacturing enterprise has some control. In the single-factor experiments, the variation of each factor is summarized in Table III. Except for the set of runs where A/F varied as in experiment 1a, A/F was set at 100%.

Changes in safety stock levels can be characterized in many ways—for example, for each part individually. We chose to multiply the safety stock levels of experiment 0 by a constant multiplier K whose value ranged from 0.5 to 2.0. Life cycle lengths were changed by using values of L to result in life cycle lengths L+6 between 6 and 24 months.

Availability Y was varied from 1 week to 12 weeks (it cannot be less than 1 week because of the 1-week shipment transit time). Y = 1 requires off-the-shelf delivery and implies a total build-to-forecast strategy. As Y increases, the production strategy shifts from build-to-forecast to build-to-order. From prior considerations, an availability Y of 18 weeks will result in on-time delivery of every order regardless of forecast quality.

While there are different ways to characterize modification of part lead times—for example, changing it for each part—we chose to change part lead times by changing the percentage of parts with lead times of 6, 10, and 14 weeks to be 100% in turn.

EOL Results. The EOL inventory graphs for A/F = 100% are summarized in Fig. 10. EOL inventory increases as safety stock increases; the results are consistent with experiment 0. When K is 0.75, we carry 12 weeks of C parts and there is no EOL RPI. When K is 1, we carry 16 weeks of C parts and end up with EOL inventory of C parts. When K is greater than 1, EOL RPI increases. When K is 2, we carry 16 weeks of B parts and EOL RPI includes both B and C parts.

Fig. 10b shows that product life length has no impact on EOL inventory. This is to be expected in the model because increasing L stretches out the middle portion of the time series graphs, and the behavior towards the end of life tends to be the same in all cases when L increases (illustrated in a future graph, Fig. 11b). For short L, the effect of the rising demand in the beginning of the life cycle affects the behavior at the end of the life cycle. Fig. 10c shows that as availability Y is shortened, EOL inventory increases, that is, quoting shorter lead times to customers exposes us to more risk of EOL inventory. This is intuitively correct; the longer the quoted availability, the longer we can afford to wait before ordering material.

Part lead time has no impact on EOL inventory when A/F = 100% (Fig. 10d).

Other Results. Fig. 11 shows the inventory measures over time as different factors are varied. Delivery performance is not shown because for A/F = 100%, delivery is always 100% on time.

Fig. 11a shows the inventory measures over time as a function of raw material safety stock multiplier K. The heights of the three initial spikes for material orders increase as K increases, directly impact RPI and on-order material, and indirectly impact on-hand and committed inventories. In general, the higher the K, the higher the inventory levels, including EOL inventory, which is the tail of the committed inventory graph. The on-order material level before the start of production increases as K increases. Keeping all the other factors constant, there is no change in backlog or delivery performance, and these are not shown in Fig. 11a.

Fig. 11b shows the inventory measures over time for varying the product life cycle by changing L from 0 to 18 months. This is one of the less interesting graphs, shown here for

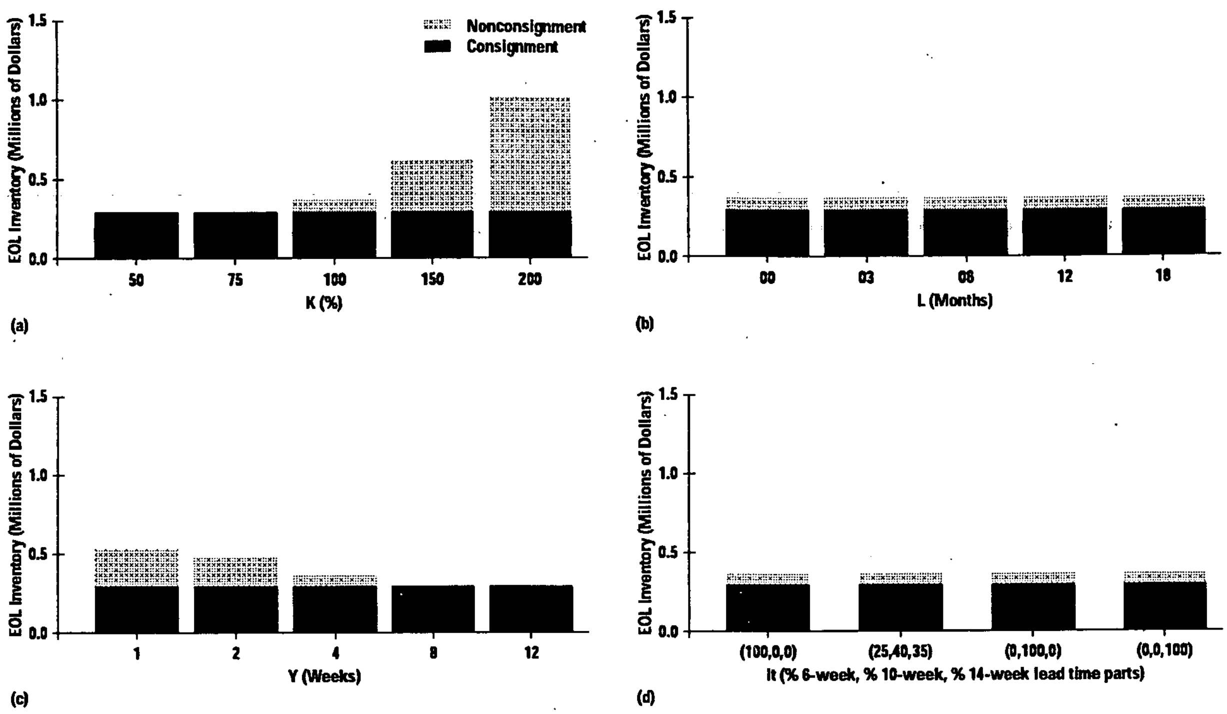


Fig. 10. EOL Inventory by single-factor changes with A/F = 100% for experiment set 1b. (a) Effect of material safety stock (5 runs). (b) Effect of product life (4 runs). (c) Effect of availability (5 runs). (d) Effect of lead time (4 runs).

completeness. EOL inventory is the same in all cases. However, because total PCFT increases, EOL inventory is a lower percentage of PCFT as L increases.

Fig. 11c shows the inventory levels over time for varying quoted availability Y. As Y increases, after the same three initial spikes, the amount of material ordered gets delayed, and the on-order material graphs get stretched to the right. The committed inventory graphs are also stretched into the future. The committed inventory is lower and the EOL inventory (tail of the committed inventory graph) tends to decrease. The delivery profiles are shifted out into the future and the backlog levels are higher.

Fig. 11d shows the time responses of the inventory metrics as part lead times vary. Notice the change in shape of the material ordered graphs. For lt = (25,40,35), there are three large and three small spikes, whereas for the other cases, there is one large spike and one small spike. As lead time increases, the material needs to be ordered earlier. On-order material increases as the lead time increases. On-hand inventory does not change. There is no impact on EOL inventory, order backlog, or on-hand inventory (RPI+WIP+FGI) as long as A/F remains constant at 100%.

Interpretation of Results. This set of results shows how each organization in the manufacturing enterprise can improve its performance metrics assuming that it relies on the forecasts given as being accurate and does not try to second-guess them. For example, if material procurement is under pressure to lower inventory levels, it would naturally try to reduce K. On the other hand, order processing and shipping would prefer to reduce Y to reduce having to deal with impatient customers.

Experiment Set 2: Dual-Factor Experiments

In this experiment set, we varied two factors in combination and attempted to observe the effects. However, instead of looking at all combinations, we looked at the impact of each of the other factors when A/F changed. This enabled us to see the effect of the controlled action in various situations of customer ordering behavior.

Results of Two-Factor Experiments. Fig. 12 summarizes the information on EOL and on-time deliveries as A/F and other factors are varied. Fig. 12a shows that as K increases, there is higher exposure to EOL inventory as A/F decreases. However, increasing K in general gives better delivery performance by shortening the average order-to-delivery time as A/F increases above 100%. Below an A/F value of 100%, K does not have an impact on the already excellent delivery performance shown by 0% late deliveries.

Fig. 12b shows that as L increases, the total shipments for a given A/F increase. For long L, the absolute volume of on-time deliveries initially increases as A/F increases. As A/F keeps on increasing past 100%, the absolute volume of on-time deliveries decreases. The average order-to-delivery time is not affected very much by L, and the EOL inventory is impacted insignificantly. The absolute amount of EOL inventory seems to depend little on L except when L is 0. For L = 0, the long lead time and high safety stock parts may actually cause most of the material for life cycle use to be ordered before the first customer order is received. The percentage of EOL writeoff decreases for a given A/F as L increases, reflecting the fact that the EOL writeoff is a smaller percentage of the total shipments as total shipments increase.

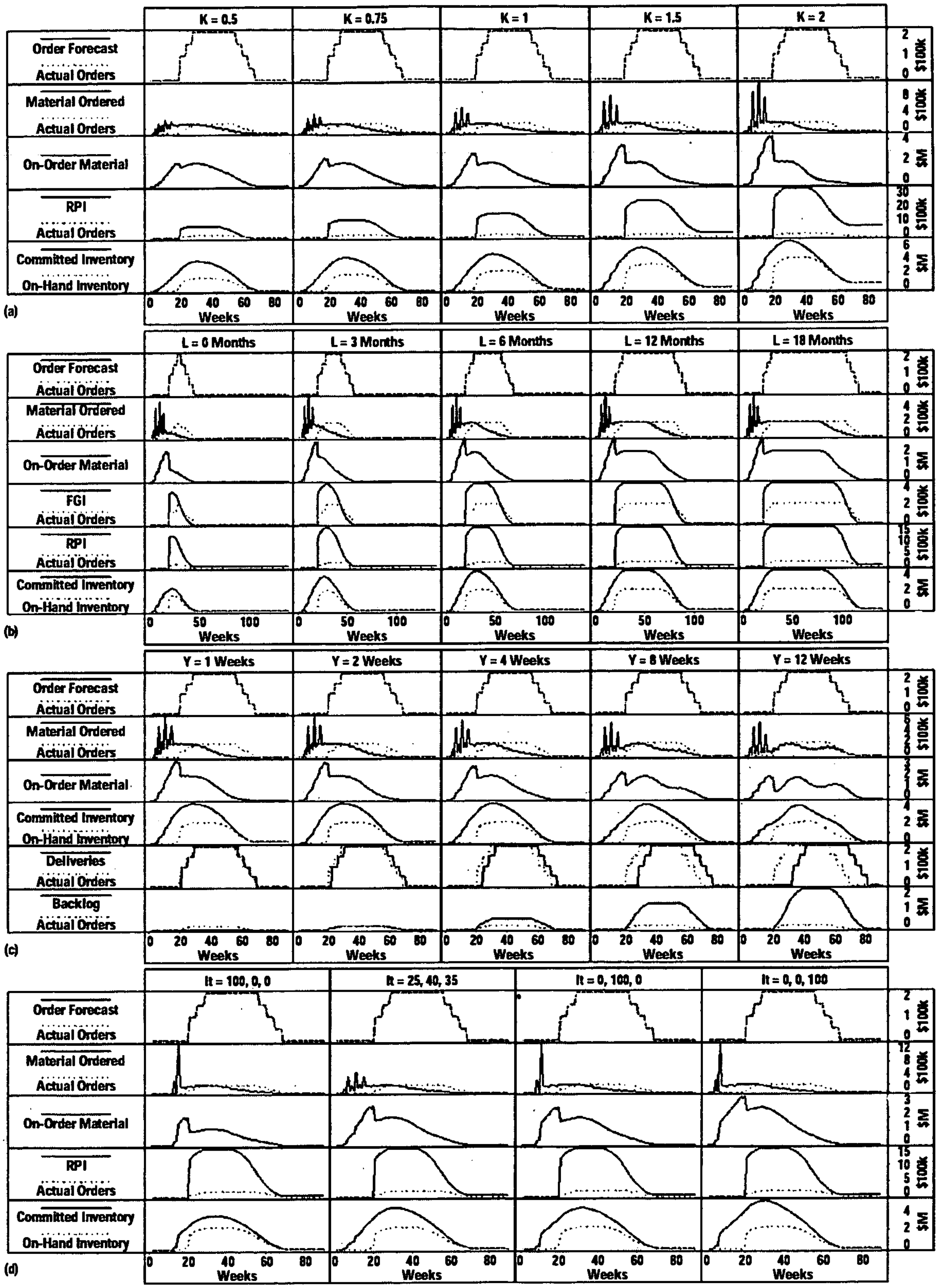


Fig. 11. Inventory measure time series for experiment set 1b: varying different factors with A/F = 100% (order forecast graph coincides with actual orders graph). (a) Varying safety stock levels (5 runs). (b) Varying life cycle (5 runs). (c) Varying availability (5 runs). (d) Varying lead time (4 runs).

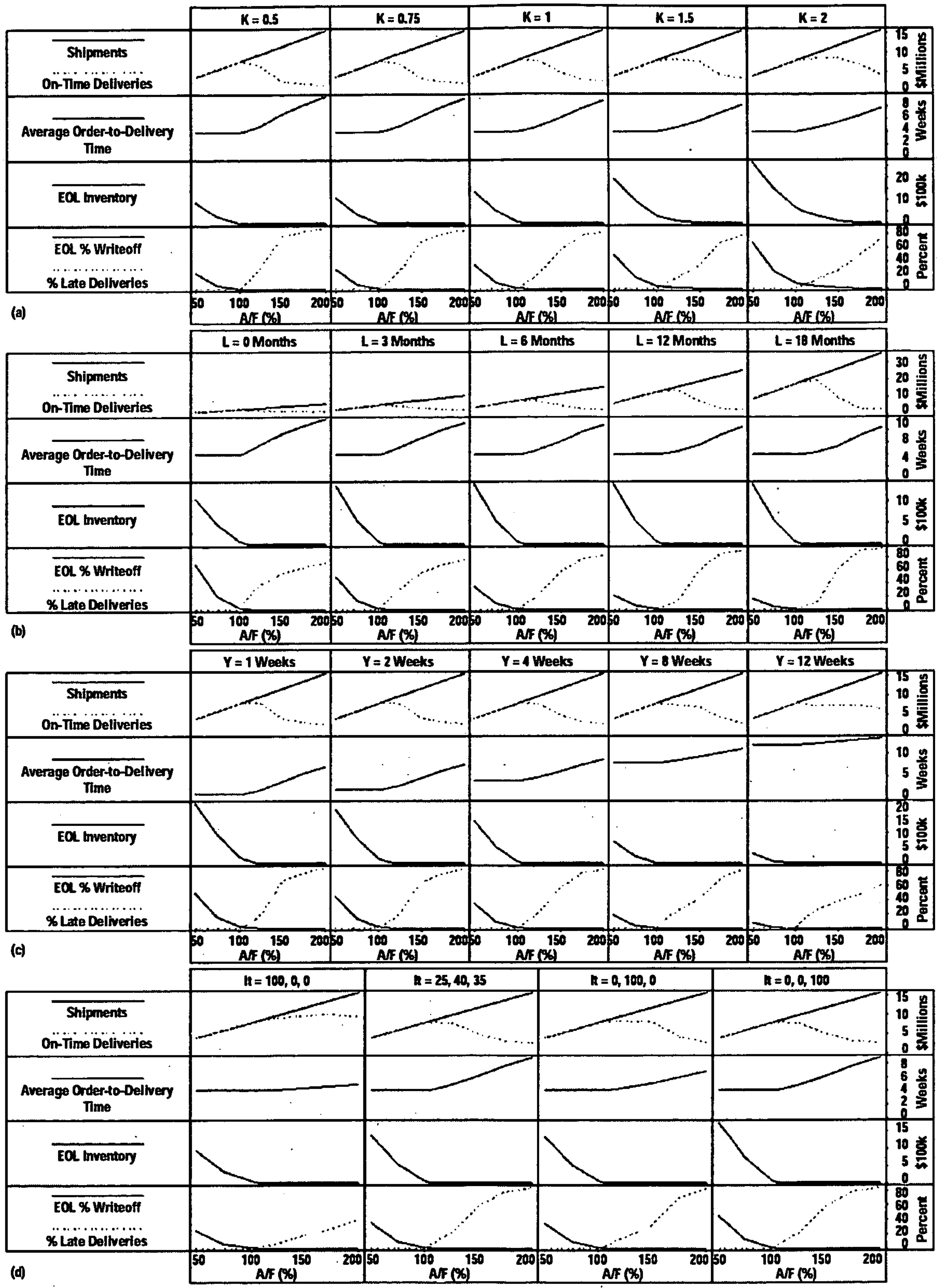


Fig. 12. EOL and shipment metrics as functions of A/F for experiment set 2 as each of the other factors is varied. (a) K varied. (b) L varied. (c) Y varied. (d) It varied.

Fig. 12c shows that increasing Y is desirable for reducing the percentage of late deliveries and reducing EOL inventory, but that the average order-to-delivery time increases, resulting in customers waiting for long periods of time, which in practice might lead to possible order cancellations. When Y = 1, the worst average order-to-delivery time is lower than the

best average order-to-delivery time when Y = 12 weeks. This is an example of a situation in which trying to reduce late deliveries by quoting a longer lead time actually leads to longer average delivery times and possibly lower customer satisfaction.

Fig. 12d shows that if all other things are kept constant, longer vendor lead time leads to poorer performance when A/F is greater than 100% and increased EOL exposure when A/F is less than 100%. For lt = (100,0,0), that is, lead time for all parts is 6 weeks, A/F has little impact on average order-to-delivery time over the given range. Furthermore, the percentage of late deliveries is generally lower than for the other values of lead time. If Y could be set to 12 weeks for the case lt = (100,0,0), no orders will ever be late, regardless of the value of A/F. Applying reasoning similar to that on page 82, the policy of waiting for customer orders to arrive before we order parts could lead to an order-to-delivery time of nine weeks, which is shorter than 12.

Observations. We have looked at the interactions of A/F with the other factors in our experiments and noticed the complexity of the interactions. The results of experiment set 2 show the impact of uncertain customer behavior on various organizations within the enterprise. In an uncertain world where A/F is outside our control, it would appear that increasing K and L, reducing lt, and increasing Y would increase on-time deliveries, which is desirable from the point of view of the manufacturing enterprise. However, increasing Y will tend to increase order-to-delivery times and backlog volumes, which could potentially lead to poorer customer satisfaction and high backlogs for order processing and shipping to deal with.

The other problem of taking these actions is that while delivery performance for the enterprise improves in general, different people and organizations are responsible for influencing and setting the values of K, Y, and lt and obtaining the reward of improved metrics. Increasing K results in better availability but increased write-off, especially if A/F is below 100%. One individual owns K, another individual owns Y, the vendors and R&D together determine lt, marketing owns L, and customers determine A/F. Any one of these can influence the other measures unilaterally, so it is necessary to coordinate the efforts of increasing some parameters and reducing others simultaneously. For example, material procurement could reduce K on the assumption that it will reduce RPI, committed inventory and EOL inventory, and this would be correct if A/F were 100%, but if A/F came in greater than 100%, the overall delivery performance would be poor. On the other hand, if R&D chose longer lead time parts because vendors demanded a premium price for short delivery times, EOL inventory would tend to be higher regardless of what value of K was chosen by material procurement. If quoted availability Y were reduced from 4 weeks to 1 week, inventory levels would tend to go up.

We could also consider the effects of the four other factors on one another, and that would give rise to another six combinations. These discussions are outside the scope of this paper.

Experiment Set F: All Factors Varied

In experiment 0, we looked at the results of one simulation run. In experiment 1, for each factor we looked at four to eleven runs. In experiment 2, we looked at 44 to 55 runs for each combination of A/F and the other factor. As we study the effects of multiple factors, the number of runs increases exponentially. Complexity increases not only in terms of number of simulation runs considered but also the way in which we analyze the data. A full factorial experiment, that is, one in which all the factors are varied in all combinations given here, requires the analysis of 5500 runs. While it is easy to specify different levels of factors, the analysis of the amount of data generated as a result of increasing the number of factors becomes intractable. For example, if all of the time series graphs of a single run were plotted on one sheet of paper each, we would have a pile of printouts eleven reams of paper thick. To do the analysis, we used a graphing technique supported in S-Plus called a design plot.†

Design Plots. Fig. 13 shows the design plots of the means of each of four different metrics at each of the levels of the five factors. The four metrics are EOL inventory, EOL inventory percentage, total on-time deliveries, and percent on-time deliveries. Each plot reflects one metric and summarizes the value of that metric for 5500 runs. The point labeled A is the mean of the EOL values of all experimental runs with A/F = 50% (mean of 500 values). A longer line indicates greater sensitivity of the metric to that factor over the range considered, all other things being equal. For example, A/F appears to have the strongest impact on EOL inventory, EOL percentage, and on-time shipments. On the other hand, the mature demand period L has a strong influence on the total dollar volume of on-time product deliveries.

An interesting point is that mean EOL and EOL percentage decrease steadily as A/F increases. On-time deliveries in dollars increases up to a point as A/F increases to 125%, but subsequently decreases (point B in Fig. 13c). The explanation is that the safety stock policy gives some protection for on-time delivery in dollars when A/F > 100%. On-time deliveries as a percentage remains at 100% for A/F ≤ 100% and subsequently decreases as A/F increases over 100% (point B′ in Fig. 13d).

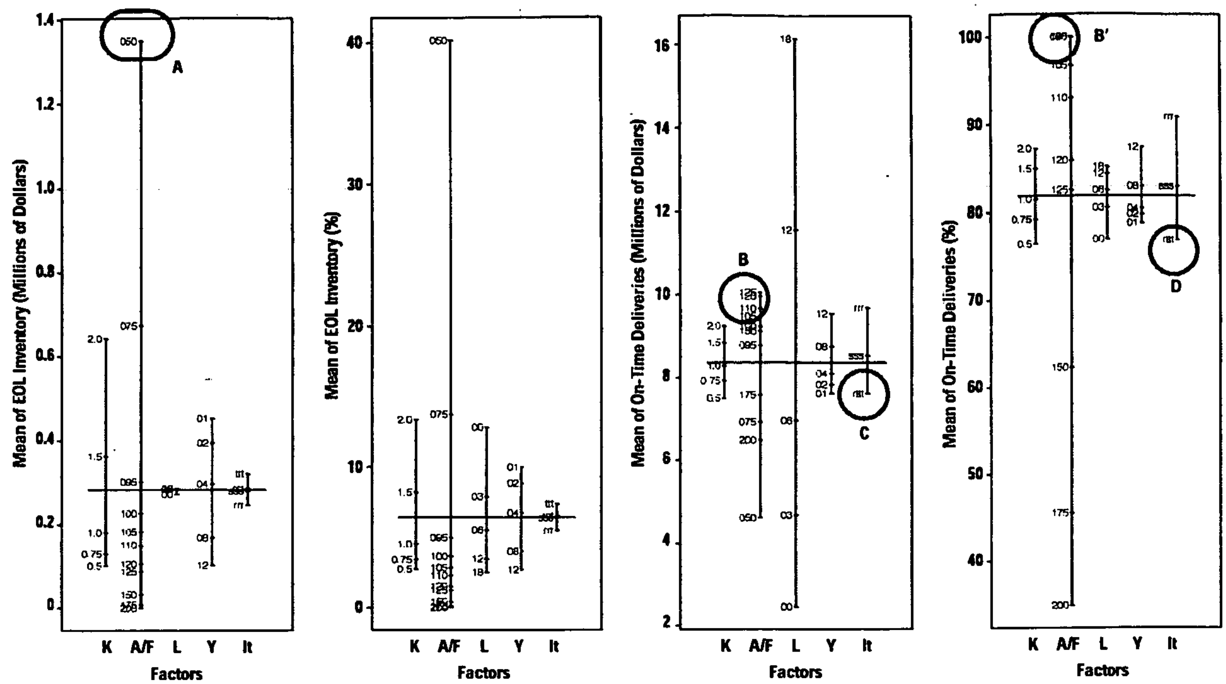
Another interesting behavior is that of the points marked C and D. The fact that the mean values of the metrics appear close together for the (25,40,35) case and the (0,0,100) case suggests that the length of the maximum lead time of parts in the bill of material has a very strong influence on on-time deliveries if all other factors are kept constant.

Further Analysis. We have barely scratched the surface of what is possible in analyzing the simulation data of this one-level bill of material, single-product situation. Further analysis and display of the variables is possible through scatter plots of pairs of variables and responses, and the use of factor plots which show greater detail. For example, further analysis could try fitting a statistical model using least sum of squares of residuals for the responses, separately and jointly. This was not done for this paper.

Experiment Set M: Multiple Product Life Cycles with Part Commonality

This set of experiments showed the impact of part commonality across multiple product life cycles. The product

† We call it a design plot because it is generated by the S-Plus function plot.design. There is no standard name of this plot. In the literature,[23] it is referred to as a "a plot of the mean response for each level of each factor."



Note: lt = (% 6-week, % 10-week, % 14-week lead time parts).
rrr indicates lt = (100,0,0)
sss indicates lt = (0,100,0)
ttt indicates lt = (0,0,100)
rst indicates lt = (25,40,35).

Fig. 13. Design plots for experiment set F: all five factors varied (5500 runs).

cycles overlapped in time, that is, one started before the preceding one finished, and we looked at a series of scenarios that differed in the values of common parts in adjacent products. These were the assumptions:

- There were four products: Adder-1, Adder-2, Adder-3, and Adder-4.
- Part commonality occurred between adjacent products only.
- Demand increased 30% for each new product.
- The unit cost of each product was 85% of the unit cost of the previous product.
- Each product life cycle was 6 months, or L = 0. This means that the complete cycle for each product is 6 months, or 24 weeks.
- There was a one-month overlap between products, that is, the first month of demand of a new product begins in the last month of demand of the previous product. This implies a total lifetime of the product family of 21 months, or 84 weeks.
- Other factors and conditions remained as in the nominal case.

Fig. 14 shows a graphical representational of the part commonality between adjacent products for the different experiments. In particular, since part commonality for experiment M-0 is 0% across adjacent products, there are no shaded areas. A fuller discussion of part commonality is given in Appendix III.

Fig. 15 shows the forecasted and actual order patterns for the four products.

Fig. 16 shows the RPI levels for parts used in the different products in Experiment M-0 (no part commonality). Consignment inventories are not shown to avoid clutter in the graphs. The WIP, FGI, products ordered, PCFT and delivery profiles are identical for all runs in experiment set M. However, each of the runs has a different profile for RPI. Note the EOL inventory of each set of parts.

Fig. 17 shows the consignment and EOL inventory levels for each run. As expected, the consignment level increases by product because the forecasted and actual orders increase by product. The consignment value for a particular product is the same across experiments. The EOL inventory for Adder-4 is the same in all the experiments. There does not appear to be any correlation between part commonality and the EOL inventory. A correlation exists between part obsolescence for a product and the EOL inventory for that product.

Fig. 18 shows the part obsolescence across products for each of the experiments. Notice how the EOL for each product in Fig. 17 is proportional to the obsolete parts for each product in Fig. 18.

Traditionally, in considering part commonality, design principles suggest using as many parts as possible from the old product. However, the results above suggest that from the point of view of EOL inventory, the amount of leftover material at the end of each product is proportional to the percentage of the part value of the obsolete parts in the old product.

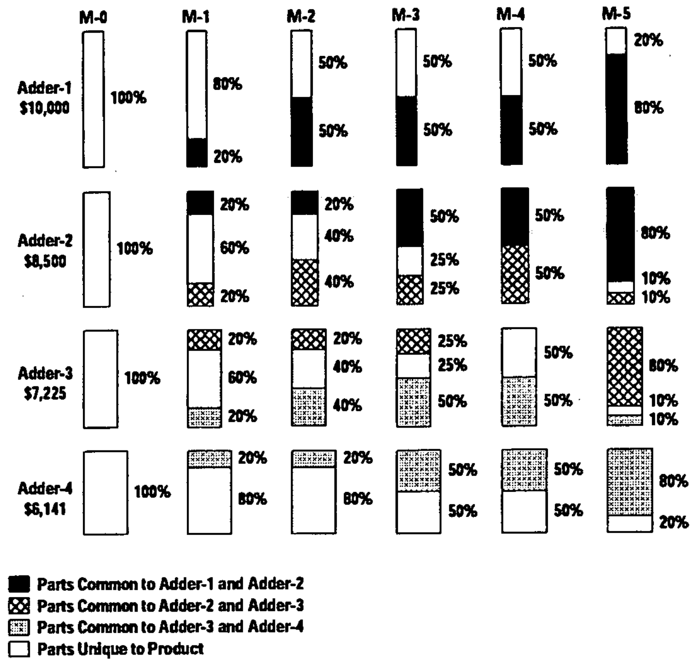


Fig. 14. Part commonality between products across experiment set M. Demand (width of bars) for each product is 30% higher than for the previous product. Unit cost (length of bars) is 85% of previous product cost.

It further suggests that the important consideration from the point of view of EOL inventory is that the percentage value of obsolete parts at the end of each product's life should be minimized.

Discussion

In this section we discuss specific results of the Simple Model, enhancements to the EMS system to do more detailed analysis, the role of the Simple Model in enterprise modeling and simulation, and optional ways of using enterprise modeling and simulation.

The major results can be summarized as follows:
- Rational material ordering and safety stock policies designed to reduce inventory to zero at the end of the product life cycle can give rise to leftover material if customers orders exactly according to forecast.
- System behavior and the impact on different metrics such as write-off, delivery times, and performance deliveries can be quantified with respect to the factors of forecast quality, safety stock levels, material lead times, product life cycles, and quoted availability individually as well as in combination.
- Forecast quality, which is influenced by the external environment, has a major effect on the metrics of interest. For example, high inventory levels may occur when actual orders come in too high or too low.
- The influence of part commonality on write-off can be quantified; this suggests an alternative way of looking at the practice of using common parts in a series of products.

What have we learned from the simulation runs on the Simple Model? We have derived a set of specific insights into system behavior under a variety of operating conditions using a methodology of generating behavior over time. We went through a large number of scenarios and showed how to gauge system behavior from the perspectives of different parties.

Interpreting the Results

The model results are sensitive to the underlying assumptions. Since we assumed the vendors always delivered on time in the simulation, the safety stocks in effect guarded only against demand uncertainties. We examined in detail the situation of order forecast bias with zero variance. This

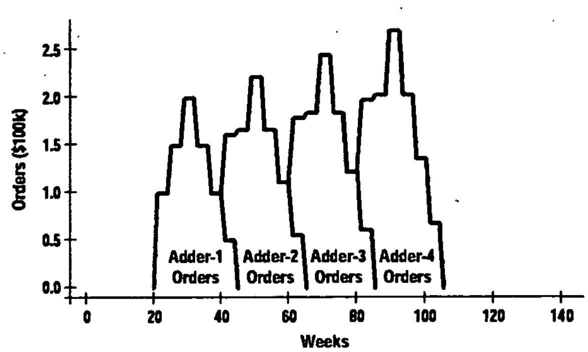


Fig. 15. Orders for different products for experiment set M.

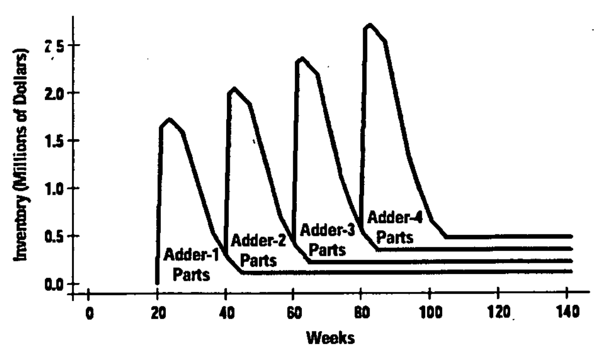


Fig. 16. RPI levels for the different parts as a function of time for experiment M-0 (no part commonality).

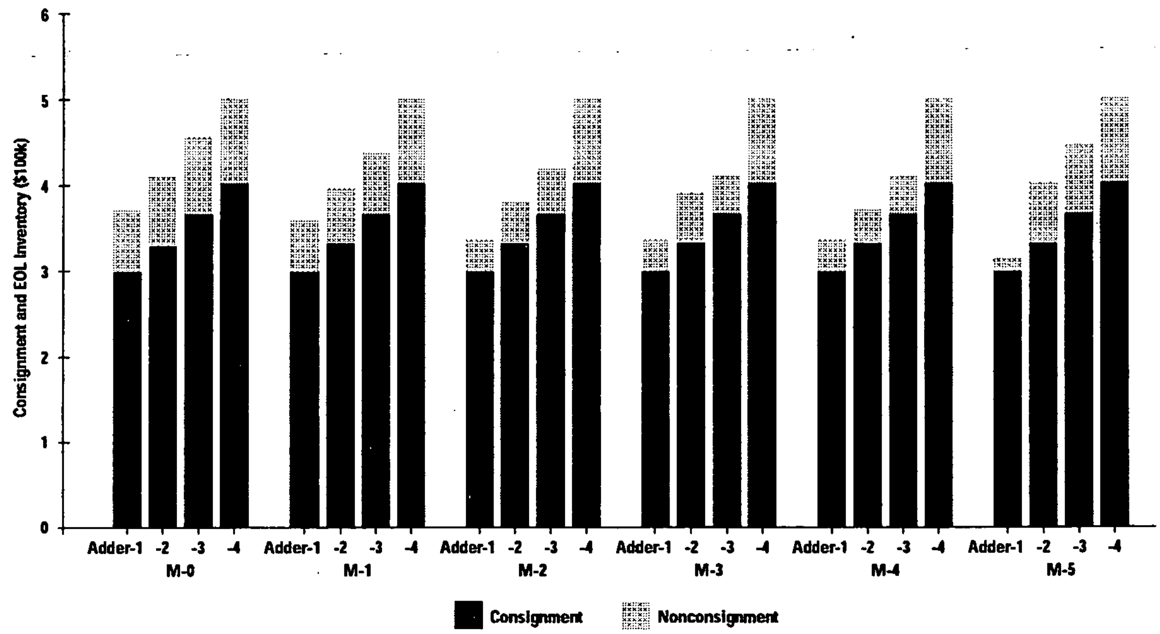


Fig. 17. Consignment and EOL inventory by product for different amounts of part commonality for experiment M-0.

is not an inherent limitation of the model, but reflects only the deterministic circumstances in which we ran the simulations. However, the results indicate that even if production and supplier lead times are completely predictable and suppliers deliver on schedule, interactions and delays within the system lead to long lead times being seen by the customers when there is underforecasting of customer orders. The manufacturing enterprise needs to take this into account and start looking elsewhere—merely making the production faster and more efficient is not sufficient.

The results so far have only scratched the surface of the analysis and interpretation possibilities. Other analysis could be done by varying ship times, FGI safety stock levels, production planning frequency, material ordering frequency,

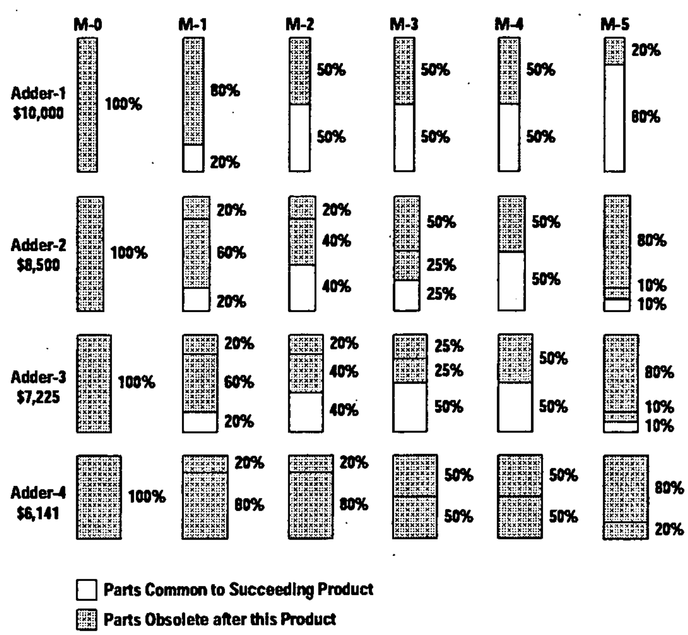


Fig. 18. Parts obsolescence between products across experiment set M.

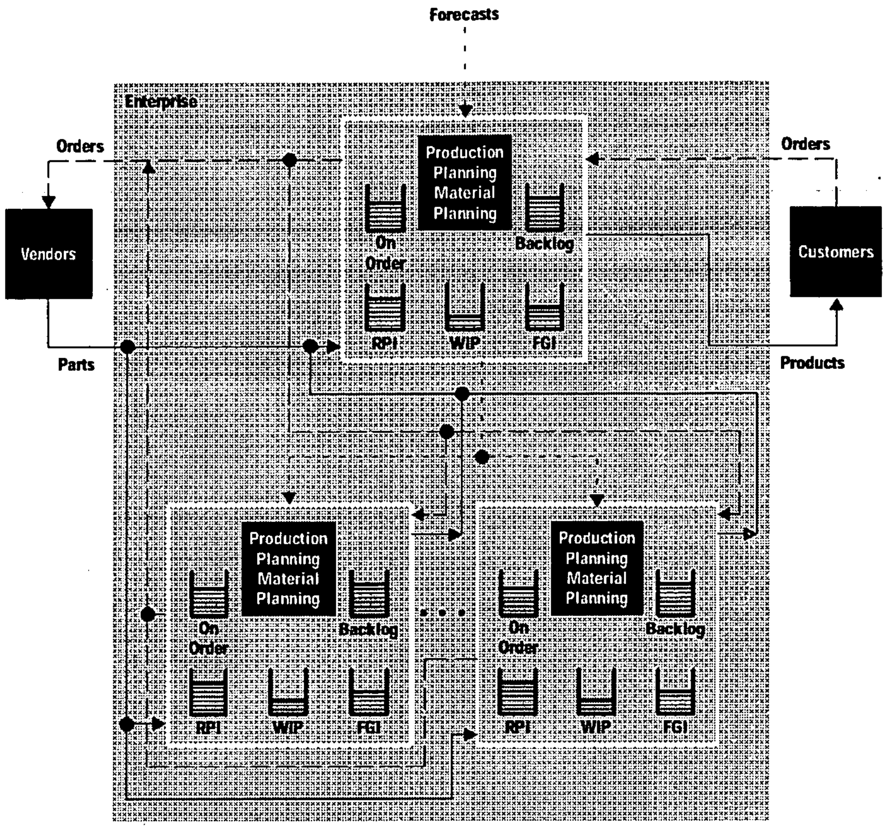


Fig. 19. Material and order flow diagram of a simple multientity distributed enterprise.

order filling policies, and uncertainty and time delays of information flow. This increases the number of runs and the quantity of data collected as well as the complexity of analysis, but would provide a richer set of relationships.

The Simple Model example may have left the reader with the impression that the current EMS system can deal with only simple or trivial cases. One goal of enterprise modeling and simulation research activities is to address successively more complex interactions and to model real-world intricacies more closely. In support of that goal, the following sections discuss subsequent and future enhancements to deal with other issues that have been raised.

Uncertainty and Variability. In the experiments described, the Simple Model was run under deterministic circumstances. Demand values and process times were constant across a particular run for convenience of understanding, and we considered uncertainty in the form of forecast biases where demands were a fixed multiple of forecasts over the period of the forecasts. Other forms of uncertainty could include the actual life cycle being different from the forecasted life cycle. Uncertainly in process times could be handled by using two values for process times: the planned process time for planning purposes and the actual time for execution. This reflects the situation when actual process times are uncertain and different from the estimated times for the process. For example, the build time for planning purposes could be two weeks, but it could turn out that the actual build time was one or three weeks.

We did not deal with variances that might occur when the total demand is forecasted accurately but the week-to-week demand fluctuates widely. Furthermore, variances of process times (e.g., delivery times from vendors and assembly times), yields (e.g., defective units), and build times for individual units were not modeled.

Dealing with variances is fairly straightforward once they are characterized. It requires using random number generators and multiple runs starting with different random number seeds—the current practice of discrete event simulation. There are three primary costs associated with this: the increase in data collection to characterize the variances of different processes, the increase in computational effort, and the increase in analysis effort. Only the data for the model needs to be changed to reflect variances. The model structure itself requires no changes.

Distribution and Multisite/Multiorganizational Interaction. The product distribution function and interaction between multiple sites were not considered in the Simple Model. Multisite and multiorganization interactions have been implemented by enclosing cloned versions of a slightly enhanced manufacturing enterprise model as shown in Fig. 19. The enhancement requires the manufacturing facility to generate and transmit its projected material requirements in addition to material orders.

Capacity and Supply Limitations. In current practice, build plans and material plans are sometimes computed ignoring production capacity and vendor limitations. In some cases, these plans are adjusted to conform to production capacity and vendor supply constraints, such as a minimum order quantity or a maximum that can be ordered in a period. In other cases, these limitations are observed at plan execution,

that is, at production, or when deliveries are not received from vendors when expected. There is no unique way of dealing with these limitations.

Implementing capacity limitations in the current Simple Model is straightforward during production. To deal with it during planning requires the inclusion of two classes of capacity constraints in the production planning algorithms: the capacity restrictions for an individual product, assembly, or subassembly as well as total capacity, and the rate at which production capacity can increase.

In reality, when prospective capacity limitations are detected, production and manufacturing line design and engineering considerations determine the rate of capacity expansion. When gross overcapacity is detected, consideration is given to reducing costs by reducing capacity. While currently the EMS system cannot model the strategic decisions of whether to expand capacity or forego extra orders, it can model the consequences of picking either of these actions.

Interaction of Multiple Products. The Simple Model assumed a single product with unconstrained production capacity. Consequently, a single unavailable part stops production of that product. Since this phenomenon also occurs with multiple products with no common parts, multiple products with no common parts can be analyzed by adding up the effects of the individual products separately. The reader familiar with linear systems will recognize this as the principle of superposition.

Adding up the results would also be valid for multiple products with common parts with no part shortages as in experiment set M. It would not be valid for multiple products with common parts, resources, and supply and production capacity limitations under scarcity conditions. When a part or resource is in short supply, decisions must be made on how to allocate the parts and resources based on some simple heuristic or optimal allocation scheme.

Multilevel Bills of Materials. The Simple Model dealt with a single-level BOM. Further expansions allow an arbitrary number of levels of BOM to be passed as data to the model. A seven-level BOM for a real product has been implemented and tested successfully. This capability to pass BOM as data allows us to make different runs with different product structures (as for example in experiment set M) without modifying the model structure.

Connection to a Mathematical Programming or Optimization Package. The Simple Model focused on applying simple algorithms for planning. The production planning and material procurement processes were initially implemented as the explicit closed-form solutions derived in Appendix I. It was realized subsequently that these algorithmic closed-form solutions were the solutions to the linear programming problem formulation. As more sophisticated planning decision techniques are proposed and studied, implementing the algorithmic solution for each new technique becomes impractical. An alternative approach is to formulate the planning process as an optimization problem and separate its solution from the formulation. This leads to concentrating on ways to better formulate the problem, leaving the solution to a separate process such as a mathematical programming package. This could provide a means of rapidly testing alternative strategies for production planning (e.g., global production planning across the entire enterprise versus local production planning at each site).

R&D, Marketing, and Cash Flow. Fig. 20 shows a proposed enterprise model at a broader scope for the next level of complexity. It generalizes Fig. 3 which focused mainly on manufacturing activities. Modeling the marketing function (and associated activities such as the forecasting process, pricing issues, and product obsolescence) could help show the impact of marketing decisions and activities on the overall system response as well as the impact of using current

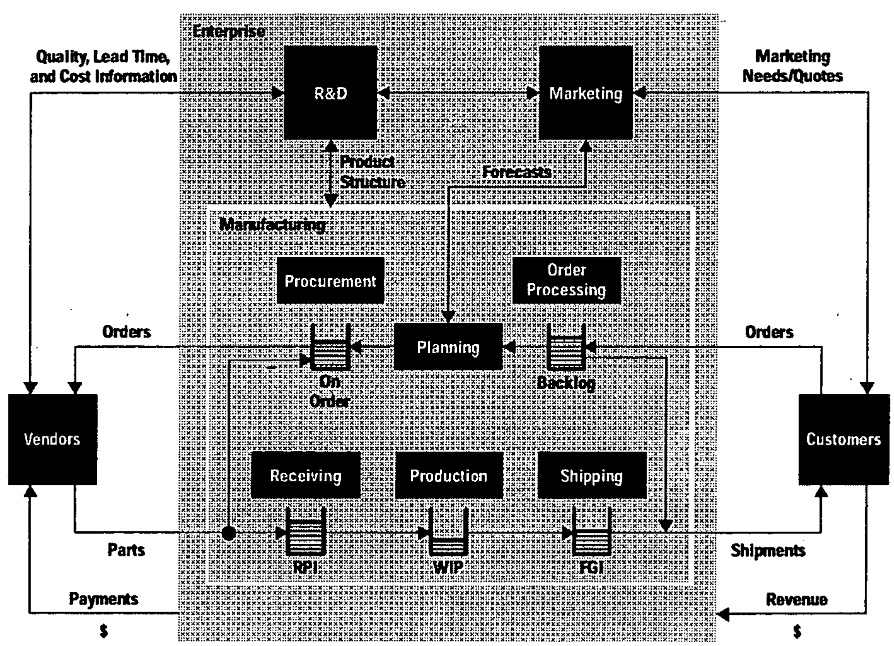


Fig. 20. Proposed enterprise modeling entities for expanded analysis.

orders to project future forecasts. Modeling the R&D function could provide insights on impacts on time to market, with product development time taken into account in addition to build time. Modeling these functions can help us deal with situations that require coordination of marketing, R&D, and manufacturing activities and can help identify the existence of leverage points for process improvement. The blocks shown in the diagram represent functions, and each could describe multiple instances of that function. For example, the block labeled manufacturing could represent multiple manufacturing sites interacting with one another.

The primary flows in the Simple Model concentrated on information (e.g., orders, forecasts, plans, and status information), material, and control (e.g., triggers that cause activities like production to start). Flows and inventory levels were converted to monetary units before being analyzed, but cash flows were not modeled explicitly.

Modeling cash flows for payments of parts, products, and process costs will provide a financial perspective. Showing projected cash flows and investments and the projected financial consequences of investment decisions will provide the stepping stones to doing discounted cash flow and net present value analysis. Modeling cash flows will also help generate pro forma financial statements to estimate revenue, cost, and income owing to different capital budgeting and allocation decisions, and provide a tool that could help address business issues. An example of such an issue is the transition from a high-margin business to a low-margin high-volume business.[24] The model may help by projecting cash requirements for investments and operations and providing estimates for return on assets during the transition.

Whither the Simple Model and the EMS System?

The Simple Model is not an end or final model; it is intermediate in a series of models that have contributed to the evolution of enterprise modeling and simulation (see page 90) and the development of the EMS system. Its simulation demonstrates the kinds of results that can be generated by enterprise modeling and simulation. Its value is in providing greater quantitative analysis where previously qualitative approaches have been adequate (see below). Its immediate subsequent application was the planning calendar model.[25,26,27]

The subsequent and future enhancements discussed make the Simple Model more complete. Some of the changes make the model larger, add detail complexity, and generate more precise results. Other changes broaden the scope of the model, and make it more representative of the other functions of the enterprise besides manufacturing; these changes require the addition of greater levels of abstraction, the ability to consolidate different points of view, and knowledge acquisition across the organization. All the changes are technically feasible and require different kinds of activities: the first set of changes requires greater emphasis on "modeling in the small," and the second set requires greater emphasis on "modeling in the large" (see discussion on page 81). Discussions based on the experience and views of some managers responsible for operations suggest that expanding the size by increasing the detail complexity, while providing greater predictability of the system, is difficult and requires a tremendous amount of investment to manage the complexity of the models and the generation and interpretation of the resulting data. Monroe[24] and Harmon[28] have individually recommended that there is greater value and potentially a far greater return on investment to be obtained by broadening the scope of future models to address and reflect business issues and concerns.

Regardless of the direction of model enhancement is the challenge of managing simulation data. The simulation runs for the experiments generated large amounts of data, and only aggregate data was collected and summarized. For instance, RPI levels for every part were generated for each week during the simulation, but the data collected was the aggregate dollar value of all the parts. The challenge became one not of collecting all data, but one of deciding ahead of time which data was interesting and not collecting that

The Simple Model: Sponsor's Perspective

As HP's Computer Systems Organization customers increasingly request delivery of complete systems with much shorter lead times, our design, manufacturing and delivery systems are being stretched beyond their performance limits.

Qualitative approaches to improvement have served us well in the past, but more quantitative analysis is needed to understand and improve the total system both from a customer and an HP perspective.

The Simple Model was conceived and developed in teamwork with HP Laboratories. We sponsored it to help learn and communicate the key drivers and characteristics of a manufacturing enterprise. The insight achieved could then be used in our order fulfillment initiative to design product, manufacturing, and delivery systems to match critical business requirements and position us to meet future customer needs effectively in the global marketplace.

Jerry Harmon
General Manager
HP Puerto Rico
Sponsor of Simple Model
for HP Computer Manufacturing

which was not; otherwise the storage requirements for storing all the generated data became significant. The data presented in the form of graphs and charts in this paper is only a small portion of the actual data collected and analyzed. A larger amount of collected data was discarded because it did not look interesting.

The sheer amount of detailed data that needs to be examined and interpreted tends to overwhelm the analyst. The analysis and interpretation of the data was very much a creative team effort requiring much discussion, and is not yet understood well enough to be automated. As we increase the number of factors, the behavior becomes more complex, and the amount of data tends to increase exponentially with the number of factors. When presented with the data in its raw form, decision makers and experts familiar with the problem issues but less familiar with modeling and simulation all have the same general reaction that it is too complex and difficult to understand. While this is a valid reaction, the reality is that the enterprise is a complex system of interacting information, material, resource, and control flows, and whether we like it or not, has complex behavior. Enterprise models as abstractions or idealizations for the real system merely reflect that complex behavior in the simulation. We can choose to ignore the complexity of the real system and use ad hoc qualitative methods to deal with the resulting behavior, or we can choose to face the complexity, understand it by selecting what we think are important factors that influence the behavior of the enterprise, and find opportunities for applying the understanding. Enterprise modeling and simulation represent one means of facing this complexity and providing an understanding of this behavior. As with most endeavors, we have found that the precursor to simplicity of expression is greater depth of understanding.

Increased technology in the hands of the modeling and simulation expert is not sufficient for providing the insight that will help make better decisions and highlight important results. Merely generating large numbers of insights and conclusions is insufficient. It requires the perspective of operations teams and decision makers to guide the direction of exploration and to emphasize the correct metrics to solve the current situation. In fact, Monroe[24] has suggested, and we in the enterprise modeling and simulation project concur, that techniques to digest and present large amounts of data rapidly and in a more easily understood fashion would be a beneficial next step and a fruitful area of research, and that joint work of a modeling expert with an operations team to further understand the issues of data reduction, interpretation, and presentation will help modeling and simulation take its rightful place as a useful tool in analyzing business decisions.

The Simple Model is a descriptive model that illustrates complex dynamic behavior of a manufacturing enterprise with low structural and detail complexity. As we have seen in this paper, its primary output is data and information on the state of the world, and it goes a great distance towards presenting observations. Unlike an optimization model, which is a prescriptive model whose solution recommends the best action under a given set of circumstances, the Simple Model does not suggest actions. It is up to the analyst or decision maker to come up with creative solutions to solve the problems highlighted by observations of the model behavior and then assess the results from a subsequent simulation run incorporating those solutions.

Prospective Applications

Let us now look at application areas for enterprise modeling and simulation. These include but are not limited to improving the performance of the current system (continuous improvement), studying the impact of reducing process times, and generating information for the enterprise, all of which are discussed below. A potentially far more powerful application is looking at new designs where the process itself is being changed (i.e., reengineering). Because of the strong current interest, large impact, and controversy surrounding reengineering, this subject is given its own discussion on page 86.

Incremental Improvements. Actions for continuous improvement can be suggested by running the nominal or baseline model and rerunning it with minor modification and changes in parameters or actions over which we have control. For example, it may not be possible to reduce all the part lead times down to six weeks, but we could certainly see the impact of reducing the value of 14-week parts in the product to determine the impact on the metrics of interest. We could look at the impact of reducing build times or FGI safety stock levels slightly to study the impact on the measures of interest. We could examine the impact of making two small changes at the same time. This application of enterprise modeling and simulation supports the process of continuous improvement by demonstrating the benefits of small changes.

Verifying Impact of Reducing Process Times. Davidow and Malone[29] talk about how short cycle times attenuate "the trumpet of doom," which is a plot of forecasting error versus time that implies that the further a person must forecast into the future, the greater the possibility of error. Rather than speculate on or guess on the impact of this trumpet of doom, enterprise modeling and simulation provide a way to

quantify the effect of reducing system cycle times. This can be accomplished by making some estimates of the amount of uncertainties within the model.

Stalk and Hout[30] suggest mapping out explicitly the major causes of problems in processes such as new product development or in operations, and comparing actual versus standard cycle times. These maps provide qualitative relationships. To the extent that processes can be mapped explicitly and quantitatively, enterprise modeling and simulation can show how the system behavior changes for a given change in the process and can verify whether modifying the component processes has the desired overall global effect.

Generating Enterprise Behavior Information. Davidow and Malone[29] identify four categories of information of use to a corporation: content, form, behavior, and action. Content information is historical in nature and reflects the experience. Form information describes shape and composition and is usually more voluminous than content information. Behavior information often begins with form information and usually requires a massive amount of computer power to predict behavior through simulation. They suggest that the final triumph of the information revolution will be the use of action information—information that instantly converts to sophisticated action. Until recently, only the most elementary category, content, has been available to business in any systematic and manageable way, and obtaining or generating the other three categories has become economically feasible only in recent years. They go on to describe how behavior information generated by computer simulation is the new paradigm for product design ranging from molecular design through automotive design to airplane design. With such behavior information design disasters of the past might be averted, and potential and unforeseen future tragedy can be replaced with a successful and predictable conclusion. With the arrival of workstations in the 1980s, it became reasonable for the computer to create realistic models and put them through their paces rather than painstakingly building prototypes and testing them under a variety of operating conditions. High-speed simulators could be built that reproduced the actual electrical characteristics of devices in different configurations.

We suggest that enterprise modeling and simulation represent an assistive and enabling technology for the design and implementation of processes of the enterprise, and that the application of such techniques to the enterprise could potentially have greater impact than product design. Furthermore, these techniques have the characteristic of converting content and form information into behavior information on which action can be taken. While the enterprise modeling and simulation process currently does not suggest actions or alternatives, it describes the behavior of the system designed with alternate processes under different operational scenarios.

Conclusions

In this paper, we outlined activities in enterprise modeling and simulation at HP Laboratories and presented in detail the results of the simulation of a simple model of a manufacturing enterprise. We have also described possible areas where enterprise modeling and simulation might be applicable, and reiterate that enterprise modeling and simulation provide a way of quantifying the impacts of proposed changes before they are implemented.

The Simple Model captures the characteristics and behavior of a manufacturing entity at a fairly high level. It shows that in the best of circumstances (e.g., customers ordering exactly according to forecast), seemingly rational operational policies can lead to end-of-life inventory. The situation only gets more complex as greater uncertainty is introduced.

Experience with using the Simple Model suggests two directions for future research in enterprise modeling and simulation. The first is to expand the scope of the Simple Model to more completely represent the functions and organizations and their interactions in the enterprise. The second is to improve the process by which the data generated by the simulation models can be understood and summarized, and the resulting information presented in a form that permits decision makers to understand more completely and to act more rapidly and with greater assurance that the desired objectives will be achieved.

Acknowledgments

The enterprise modeling and simulation system was developed by the team of Bob Ritter, Bob Joy, and the author, all of whom are affiliated with Hewlett-Packard Laboratories. The Simple Model was proposed by Jerry Harmon of HP Puerto Rico and developed by Shailendra Jain and the author in two parallel efforts that served to verify the results using two different approaches. Jerry Harmon, Bob Ritter, Shailendra Jain, and Paul Williams of Hewlett-Packard Laboratories and the author were involved in interpreting the results of the Simple Model at various times. Others have provided fresh insight during the course of discussions. The author would like to thank his team members, colleagues, and the many reviewers for their helpful comments and assistance in the preparation of this document.

References

1. A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis, Second Edition,* McGraw-Hill, Inc., 1991, p. 1.
2. A.A.B. Pritsker, *Introduction to Simulation and SLAM II, Third Edition,* Systems Publishing Corp., 1986.
3. R. McHaney, *Computer Simulation—A Practical Perspective,* Academic Press, Inc., 1991.
4. A.M. Law and M.G. McComas, "Secrets of Successful Simulation Studies," *Proceedings of the Winter Simulation Conference 1991,* Society for Computer Simulation, pp. 21-27.
5. F.E. Cellier, *Continuous System Modeling,* Springer-Verlag, Inc., 1991.
6. P.M. Senge, *The Fifth Discipline,* Doubleday/Currency, 1990.
7. L. Marran, M.S. Fadali, and E. Tacker, "A New Modeling Methodology for Large-Scale Systems," *Proceedings of the International Conference on Systems, Man, and Cybernetics,* IEEE, May 1989, pp. 989-990.
8. *Structured Methods—An Overview for Engineers and Managers,* Hewlett-Packard Corporate Engineering, 1988, pp. 77-90.
9. M.S. Mujtaba, "Simulation Modelling of a Manufacturing Enterprise with Complex Material, Information, and Control Flows," *International Journal of Computer Integrated Manufacturing,* Vol. 7, no. 1, 1994, pp. 29-46.
10. B.P. Zeigler, *Multifaceted Modeling and Discrete Event Simulation,* Academic Press Inc., 1984.

11. V.B. Norman, et al, "Simulation Practices in Manufacturing," *Proceedings of the Winter Simulation Conference 1992,* pp. 1004-1010.

12. M. Fox, "The TOVE Project—Towards a Common Sense Model of the Enterprise," *Proceedings of the International Conference on Object-Oriented Manufacturing Systems (ICOOMS),* May 1992, pp. 176-181.

13. H.R. Jorysz and F.B. Vernadat, "CIM-OSA Part 1: Total Enterprise Modelling and Function View," *International Journal of Computer Integrated Manufacturing,* Vol. 3, nos. 3 and 4, Fall 1990, pp. 144-156.

14. H.R. Jorysz and F.B. Vernadat, "CIM-OSA Part 2: Information View," *International Journal of Computer Integrated Manufacturing,* Vol. 3, nos. 3 and 4, Fall 1990, pp. 157-167.

15. A. Pardasani and A. Chan, "Enterprise Model: A Decision-Support Tool for Computer Integrated Manufacturing," *Proceedings of the International Conference on Object-Oriented Manufacturing Systems (ICOOMS),* May 1992, pp. 182-187.

16. J. Harmon, *personal communication.*

17. R. Norton, "A New Tool to Help Managers," *Fortune,* May 30, 1994, pp. 135-140.

18. G.L. Steele, *Common Lisp—The Language, Second Edition,* Digital Press, 1990.

19. *Allegro CL Common Lisp User Guide Volumes 1 and 2 Version 4.2,* Franz Inc., January 1994.

20. *Lucid Common Lisp/HP Manuals and User Guide,* Lucid Inc., June 1990.

21. *LispWorks Manuals Edition 3.2,* Harlequin Ltd., March 1994.

22. M.S. Mujtaba, *Formulation of the Order-to-Ship Process Simulation Model,* HP Laboratories Technical Report #HPL-92-135, December 1992.

23. J.M. Chambers and T.J. Hastie, *Statistical Models in S,* Wadsworth & Brooks/Cole Advanced Books & Software, 1992.

24. J. Monroe, sponsor of planning calendar model,[25,26,27] leader of HP Computer Systems Organization planning program 1992-1993, *telephone communication,* September 12, 1994.

25. C.M. Kozierok, *Analysis of Inventory and Customer Service Performance Using a Simple Manufacturing Model,* Master of Science Thesis for Leaders for Manufacturing (LFM) Program, Massachusetts Institute of Technology, May 1993.

26. K. Oliver, *Simple Model Report,* distributed by email on January 12, 1993.

27. M.S. Mujtaba and R. Ritter, *Enterprise Modeling System: Inventory Exposure and Delivery Performance,* HP Laboratories Technical Report #HPL-94-89, October 1994.

28. J. Harmon, sponsor and proponent of the Simple Model, leader of HP Computer Manufacturing forecasting and planning redesign team 1991-1992, currently General Manager, HP Puerto Rico, *telephone communication,* September 1994.

29. W.H. Davidow and M.S. Malone, *The Virtual Corporation,* HarperCollins Publishers, Inc., 1992.

30. G. Stalk and T.M. Hout, "Competing Against Time," *The Free Press,* A Division of Macmillan, Inc., 1990.

Appendix I: Mathematics of Production and Material Planning for the Simple Model

I-1 The Planning Function

The planning function is actually an analytic model embedded within a discrete event simulation model. The fundamental principle on which the production and material planning algorithms are based is the conservation of mass, that is, consumption cannot be higher than the total supply available. The order in which the build plan computation is done is the reverse of the order in which subassemblies are built and products are shipped (i.e., from shipment to product build to part order). For ease of explanation, the current week is considered to be week 0. This derivation emphasizes clarity of explanation rather than rigorous detail.

There are three sets of decision variables to be determined for each week: *s(t)*, the shipment plan, *b(t)*, the build plan, and $m_j(t)$, the material ordering plan. These are shown in italics.

Before we get into the mathematical formulation, let us first look at the process of computation. Fig. 1 illustrates how the production and material planning algorithms work in this model. The computational process is described in the following order:

- I-2 describes the notation shown in Fig. 1.
- I-3 describes the safety stock computation.
- I-4 describes the initial conditions for computation.
- I-5 describes the computation of the shipment plan.
- I-6 describes the computation of the build plan.
- I-7 describes computation of the number of units started this week.
- I-8 describes the computation of the material consumption and material ordering plans.
- I-9 describes the actual material ordered this week.
- I-10 describes the computation of the number of weeks for each of the plans.

I-2 Notation

- n, s, t = indexes for week number (current week = 0)
- f(t) = Current forecast of product orders for week t, t = 0, 1, ..., N_f
- F(t) = FGI at end of week t
- W(t) = WIP at end of week t
- B(t) = Backlog units at end of week t
- B(t,s) = Backlog units at end of week t having shipment dates in week s
- *s(t)* = Planned shipments during week t
- *b(t)* = Units planned to be started during week t
- B = Build time in number of weeks
- Y = Quoted availability in number of weeks
- S = Shipment or transit time
- j = Index relating to part
- Q_j = Quantity of part j per unit of product
- $q_j(t)$ = Planned consumption of part j during week t

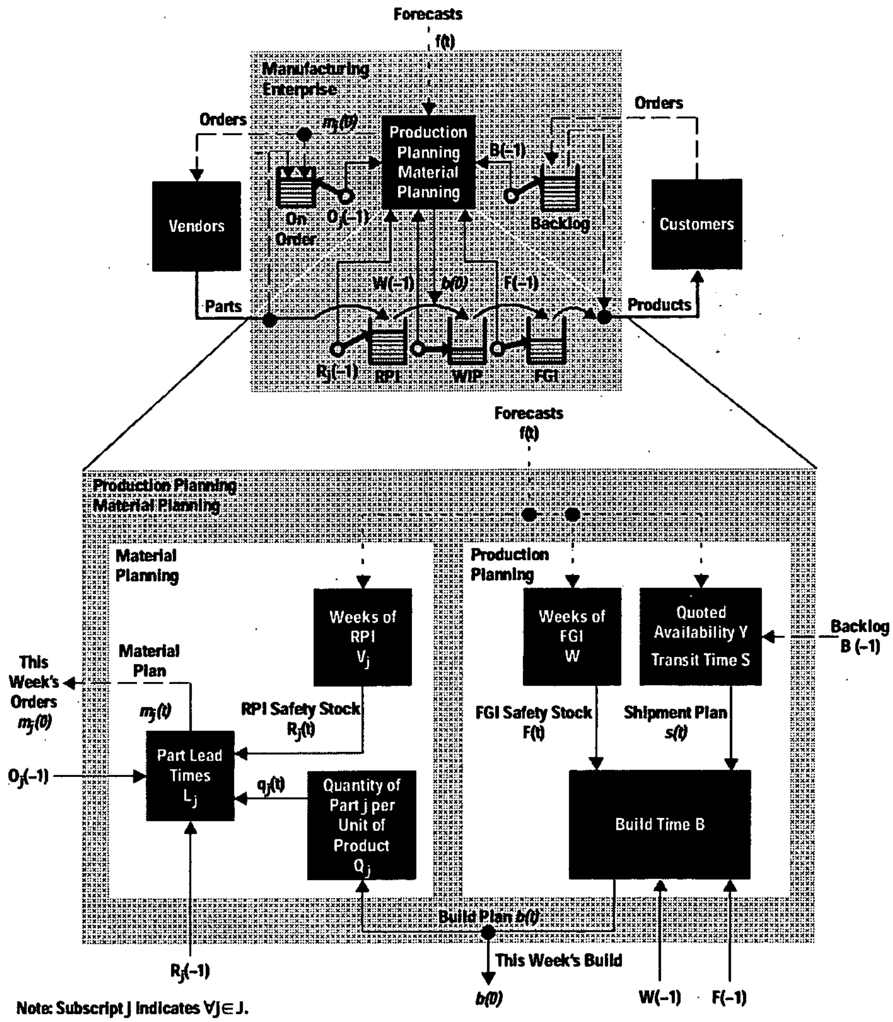


Fig. 1. Notation and production/material planning. The shipment plan is computed from the backlog, forecasts, quoted availability, and transit time. The build plan is computed from the shipment plan, the build time, WIP, FGI, and FGI safety stock. The actual build is computed from the build plan and the material availability. The material consumption plan is computed from the build plan and the bill of materials. The material ordering plan is computed from RPI, RPI safety stock, the material consumption plan, on-order material, and lead time.

- $m_j(t)$ = Planned quantity of material j to be ordered during week t, t = 0, 1, ..., N_j, $j \in J$
- $R_j(t)$ = RPI of part j at the end of week t
- $r_j(t)$ = Units of part j received during week t
- $O_j(t)$ = Units of part j on order at the end of week t
- L_j = Vendor lead time for part j
- J = Set of parts that go into the product
- w = FGI safety stock in weeks of demand
- V_j = RPI safety stock of part j in weeks of demand
- N_s = Last week for computing shipment plan
- N_b = Last week for computing build plan
- N_f = Last week used for forecasts
- N_j = Last week for computing material order for part j.

Since the current week is 0, the values of these variables represent actual values for weeks before 0, and the values are computed, set, or derived for weeks 0 and later. In particular, the values of variables at the end of week –1 represent the current values of those variables, as described in I-4. All numerical quantities except time indexes are zero or positive.

I-3 Safety Stock Computation

Safety stock is expressed in number of weeks of 13-week leading average forecast. The 13-week leading average forecast at the end of week t is defined as:

$$\overline{f(t)} = \frac{1}{13}\sum_{i=1}^{13} f(t+i) \qquad (1)$$

The target FGI safety stock at the end of week t is w weeks and the target RPI safety stock at the end of week t for part j is V_j weeks. The expressions for these quantities are:

$$F(t) = w\overline{f(t)} \qquad (2)$$

$$R_j(t) = V_j Q_j \overline{f(t)} \qquad (3)$$

I-4 Initial Conditions

- F(–1) = Actual FGI at the end of the previous week, that is, current FGI
- W(–1) = Actual WIP at the end of the previous week, that is, current WIP
- $O_j(-1)$ = Actual part j on order at the end of the previous week, that is, current on-order material
- $R_j(-1)$ = Actual RPI for part j at the end of the previous week, that is, current RPI for part j.
- B(–1) = Order backlog in units at the end of the previous week, that is, current backlog:

$$B(-1) = \sum_{s \in (\text{all shipment dates in current backlog})} B(-1, s) \qquad (4)$$

- B(–1,s) = Component of current backlog with shipment date in week s.

I-5 Shipment Plan

The shipment plan indicates prospective shipments during the current and future weeks. It is computed on the assumption that customer orders are not shipped before they are due, but are shipped in time to satisfy the quoted availability requirements. This implies that for any week, the orders planned to be shipped are those that are already late (i.e., should have been shipped in an earlier week) and those that must be shipped to be delivered on time. Notice that in computing the shipping plan, we do not take into account the amount of inventory on hand or in process. This is representative of the way shipment plans are computed and then subsequently checked against reality.

Put another way, this can be expressed as planning to ship the minimum quantity in each week that will satisfy the quoted availability criteria. The problem can be formulated as shown in the set of equations below, which indicate that we are attempting to minimize shipments in the current week, current plus next week, current plus next 2 weeks, and so on such that the total shipments in those weeks is greater than the current existing backlog whose shipment date is already past or in those weeks, plus the forecasted orders whose desired shipment dates lie in those weeks.

Minimize $s(n)$, n = 0,1,...,N_s

such that $\sum_{t=0}^{n} s(t) \geq \sum_{t \in (i|i \leq n)} B(-1, t) + \sum_{t=0}^{n-(Y-S)} f(t)$

and $s(n) \geq 0$.

These equations define a series of (N_s + 1) linear programming problems. However, this formulation will always return a set of feasible solutions, and the optimal feasible solutions can be expressed in closed form as follows:

$$s(n) = \begin{cases} \sum_{s \in (i|i \leq 0)} B(-1, s) & \text{for } n = 0 \\ B(-1, n) & \text{for } 0 < n < Y - S \\ f(n - (Y - S)) & \text{for } n \geq Y - S. \end{cases} \qquad (5)$$

The term (Y – S) is the difference between the quoted availability and the transit time (i.e., the order-to-ship time to achieve on-time delivery), and indicates the time in the future after which shipments depend solely on forecasts.

I-6 Build Plan

The build plan, which indicates how many units are to be started in the current week 0 and succeeding weeks, is based on the assumption that the FGI levels at the end of weeks 0,1,...,B–1 have already been determined by the current FGI, WIP, and shipments preceding week 0. It further assumes that we might be able to control FGI at the end of week B or later by deciding how many units we start this week and future weeks, that is, by controlling $b(0), b(1), \ldots, b(n)$. We want to keep the $b(n)$ as low as possible but greater than or equal to 0, such that the total planned build during weeks 0 through n must be greater than or equal to shipments during weeks 0 through B+n plus FGI at the end of week B+n minus current FGI and WIP. The complete formulation is as follows:

Minimize $b(n)$, n = 0,1,...,N_b

such that $\sum_{t=0}^{n} b(t) \geq \sum_{t=0}^{B+n} s(t) + F(B+n) - F(-1) - W(-1)$

and $b(n) \geq 0$.

Again, the above is a series of (N_b+1) linear programming problems, with optimal feasible solutions that are expressed in closed form as follows:

$$b(n) = \max\left\{0,\ F(B+n) + \sum_{t=0}^{B+n} s(t) - F(-1) - W(-1) - \sum_{t=0}^{n-1} b(t)\right\}, \qquad (6)$$

for n = 0, 1, ..., N_b.

To summarize the above, the current build plan should look as follows:

Week:	0	1	2	...	n
Planned Build:	*b(0)*	*b(1)*	*b(2)*	...	*b(n).*

I-7 Actual Units Started

The actual units started this week, b_0, will be $b(0)$ if there is sufficient material. If there is insufficient material the actual units started is the maximum possible with the available material, or:

Maximize b_0

such that $Q_j b_0 \leq R_j(-1) + r_j(0),\ \forall j \in J$

and $0 \leq b_0 \leq b(0)$,

for which the closed form solution is:

$$b_0 = \min\left\{ b(0), \min_{j \in J}\left(\frac{R_j(-1) + r_j(0)}{Q_j}\right)\right\}. \qquad (7)$$

I-8 Material Requirement Analysis

If the lead time for a part j is L_j weeks, the RPI level for part j at the end of weeks 0,1, ... L_j–1 has been determined by material on hand, material on order, and projected use. We could control RPI for part j at the end of week L_j or later by deciding how much of part j we order in this week and subsequent weeks. The estimated material consumption during a week is the quantity of the material for the build for that week, that is:

$$q_j(t) = Q_j b(t). \tag{8}$$

The material ordered during weeks 0 through n must be greater than or equal to the material consumed during weeks 0 through L_j+n plus the desired safety stock at the end of week L_j+n minus the current on-hand material and the current on-order material. This can be expressed mathematically as follows:

Minimize $m_j(n)$, n = 0,1,...,N_j, $j \in J$

$$\text{such that } \sum_{t=0}^{n} m_j(t) \geq \sum_{t=0}^{L_j+n} q_j(t) + R_j\left(L_j + n\right) - R_j(-1) - O_j(-1)$$

and $m_j(n) \geq 0$.

After substituting equation 8, this becomes a series of linear programming formulations for which the closed form solution is:

$$m_j(n) = \max\left\{\begin{matrix} 0 \\ \left(Q_j \sum_{t=0}^{L_j+n} b(t) + R_j\left(L_j + n\right) - R_j(-1) - O_j(-1) - \sum_{t=0}^{n-1} m_j(t) \right) \end{matrix}\right. \tag{9}$$

for n = 0, 1, ..., N_j, $j \in J$.

The current material ordering plan is shown by the following table.

	Week				
	0	**1**	**2**	...	**n**
Material 1	$m_1(0)$	$m_1(1)$	$m_1(2)$	...	$m_1(n)$
Material 2	$m_2(0)$	$m_2(1)$	$m_2(2)$	...	$m_2(n)$
...	...	...	...	...	...
Material j	$m_j(0)$	$m_j(1)$	$m_j(2)$	...	$m_j(n)$
...	...	...	...	...	...

I-9 Actual Material Ordered

Given the table above, the actual material ordered in this week must be $m_j(0)$, $\forall j \in J$.

I-10 Determination of the Required Number of Weeks

Since we want to compute the material procurement plan for material j for periods 0 through N_j, we need to make sure we have values of the forecasts, shipment plan, and build plan far enough in the future to allow us to do so. This section shows how many periods of those plans we need to compute.

In 10 through 16 below, "$m_j(n)$ requires x(n)" should be read as, "Computing $m_j(n)$ requires values of x(0), x(1), ..., x(n)." Thus 10 should be read as, "Computing $m_j(N_j)$ requires the values of $R_j(0)$, $R_j(1)$, ..., $R_j(L_j+N_j)$."

From 9,

$$m_j(N_j) \text{ requires } R_j(L_j + N_j) \tag{10}$$

$$\text{and } m_j(N_j) \text{ requires } b(L_j + N_j). \tag{11}$$

From 10, 3, and 1,

$$m_j(N_j) \text{ requires } f(L_j + N_j + 13). \tag{12}$$

From 11 and 6,

$$m_j(N_j) \text{ requires } F(B + L_j + N_j) \tag{13}$$

$$\text{and } m_j(N_j) \text{ requires } s(B + L_j + N_j). \tag{14}$$

From13, 2, and 1,

$$m_j(N_j) \text{ requires } f(B + L_j + N_j + 13). \tag{15}$$

From 14, 5, and 1,

$$m_j(N_j) \text{ requires } f(B + L_j + N_j - (Y - S)). \tag{16}$$

Computation of N_b. From 11,

$$N_b = \max_{j \in J}\left\{L_j + N_j\right\}. \tag{17}$$

Computation of N_s. From 14,

$$N_s = \max_{j \in J}\left\{B + L_j + N_j\right\}. \tag{18}$$

Computation of N_f. From 12, 15, and 16,

$$N_f = \max_{j \in J}\left\{\begin{matrix} L_j + N_j + 13 \\ B + L_j + N_j + 13 \\ B + L_j + N_j - (Y - S) \end{matrix}\right. \tag{19}$$

Since $B \geq 0$, $(Y - S) \geq 0$, the middle expression dominates, and 19 reduces to:

$$N_f = \max_{j \in J}\left\{B + L_j + N_j + 13\right\}. \tag{20}$$

Appendix II: Weekly Event Sequence

In the following table, periodically scheduled events are shown in sequence.

Event Time	Event Frequency	Initiators	Event Description
Monday 1:00	Weekly	Customers	Generate and send orders; these orders are received by the Adder factory at 9:30:00 the same day.
Monday 8:00	Weekly	Factory	Completes computing FGI safety stock for future weeks. Completes computing shipment plan and build plans.
Monday 9:00	Weekly	Factory	Completes computing material requirements plan. Completes computing material procurements plan.
Monday 10:00	Weekly	Factory	Generates current week's material orders. Material orders arrive at the vendors instantaneously.
Monday 10:00:01	Weekly	Vendors	Finish filling and shipping orders due this week. Shipments arrive at the factory instantaneously.
Monday 10:30	Weekly	Factory	Begins current week's production. Completes production started two weeks ago.
Friday 16:30	Weekly	Factory	Completes filling and shipping orders for the week.
Friday 23:58	Weekly	Simulation Executive	Records values of all the state variables.

Appendix III: Details of Part Commonality Experiments

The following table shows the definitions used to describe part commonality. MC stands for material cost, with uppercase denoting dollar values and lowercase denoting percentage values. m represents the set of material.

	Set of Material	Value of Material	Percentage Value
Common to products i and i−1	$m_{i,i-1}$	$MC_{i,i-1}$	$mc_{i,i-1} = \frac{MC_{i,i-1}}{MC_i} \times 100$
Unique to product i	$m_{i,i}$	$MC_{i,i}$	$mc_{i,i} = \frac{MC_{i,i}}{MC_i} \times 100$
Common to products i and i+1	$m_{i,i+1}$	$MC_{i,i+1}$	$mc_{i,i+1} = \frac{MC_{i,i+1}}{MC_i} \times 100$

Commonality occurs only between adjacent products. This implies that a part can be used in at most two products.

Each of the $MC_{i,j}$ is further broken up into class A, B, and C parts with relative values 50, 30, and 20 percent. Each of these classes is made up of 6, 10, and 14 week lead times with relative values 25, 40, and 35 percent. (See Table I on page 83.)

At the end of the product i life cycle, obsolete inventory (if any) should come only from parts in sets $m_{i,i}$ and $m_{i,i-1}$. Any leftover parts from $m_{i,i+1}$ can be used in product i+1. This implies that $mc_{i,i-1}$ and $mc_{i,i}$ impact the obsolete inventory at the end of the product life cycle for product i.

The values shown in the following table should be derived from the real bill of materials. For our experiments, we reverse the process, that is, we generate a bill of materials from the table, which was generated heuristically from the experimental scenarios, with the following constraints on the values of mc:

- For each i and j, $mc_{i,j}$ must be greater than or equal to 0 and less than or equal to 100.
- For each i, the sum of $mc_{i,j}$ over all j must be 100.
- In each experiment, if any $mc_{i,i+1}$ is zero, then $mc_{i+1,i}$ must also be zero.

Description of Experimental Scenarios

Run M-0: no part commonality at all between adjacent products.

Run M-1: 20% part commonality between adjacent products. The parts common to products i and i+1 make up 20% of the part values of both products. This may happen by a reduction in either part quantity or part cost, but the reason is not reflected in the dollar value of leftover inventory or material.

Run M-2: 20% part commonality when moving to a new product. The parts common to products i−1 and i make up 20% of the part value of product i; the rest of the value of product i is split equally between the parts unique to product i and those common to products i and i+1. Since product Adder-1 has no prior product, the value is split equally between unique parts and parts common to Adder-1 and Adder-2. 20% of the value of Adder-2 is made up of parts common to Adder-1 and Adder-2; the remaining 80% is split equally between unique parts and parts common to Adder-2 and Adder-3. 20% of the value of product Adder-4 is made up of parts common to Adder-3 and Adder-4; the balance of the value is unique parts since there are no succeeding products.

Run M-3: 50% and 25% part commonality between alternate products. There is 50% part commonality between products Adder-1 and Adder-2 and between Adder-3 and Adder-4; there is 25% part commonality between Adder-2 and Adder-3.

Run M-4: 50% part commonality between adjacent products; no unique parts in Adder-2 and Adder-3; 50% unique parts in Adder-1 and Adder-4.

Run M-5: 80% part commonality between succeeding products.

Part Commonality Data (%) for Multiple Product Crossover

i	Product	Demand (units)	Product Cost ($)	Common Parts (%)	Experiment Run M-0	M-1	M-2	M-3	M-4	M-5
1	Adder-1	V	10,000	$mc_{1,1}$	100	80	50	50	50	20
				$mc_{1,2}$	0	20	50	50	50	80
2	Adder-2	1.3V	0.85 × 10,000	$mc_{2,1}$	0	20	20	50	50	80
				$mc_{2,2}$	100	60	40	25	0	10
				$mc_{2,3}$	0	20	20	25	50	10
3	Adder-3	1.3 × 1.3V	0.85 × 0.85 × 10,000	$mc_{3,2}$	0	20	20	25	50	80
				$mc_{3,3}$	100	60	40	25	0	10
				$mc_{3,4}$	0	20	40	50	50	10
4	Adder-4	1.3 × 1.3 × 1.3V	0.85 × 0.85 × 0.85 × 10,000	$mc_{4,3}$	0	20	20	50	50	80
				$mc_{4,4}$	100	80	80	50	50	20

Appendix IV: Details of Explanations for Experiments 0 and 1a

IV-1 Estimated Financial Impact Based on Theoretical Considerations for Experiment 0

The impact of product Adder on the financial situation of the enterprise, as explained on page 89, is:

- Total PCFT = $7,800,000
- Mature volume = MV = mature PCFT = $800,000/month or $200,000/week
- Consignment inventory = $300,000.

IV-2 Mature Demand Week Considerations for Experiment 0

RPI Material to Support Mature Demand

		Class A	Class B	Class C	All Classes
①	Percentage of Part Value in Product	50%	30%	20%	100%
②	Weekly Use during Mature Demand ①×MV	$100k	$60k	$40k	$200k
③	RPI Safety Stock in Weeks	4	8	16	N/A
④	RPI in $: ③×MV	$400k	$480k	$640k	$1520k
⑤	RPI in Weeks of MV ④÷MV	2	2.4	3.2	7.6

On-Order Material to Support Mature Demand

①	Lead Time	6 weeks	10 weeks	14 weeks	All Parts
②	Percentage of Part Value in Product	25%	40%	35%	100%
③	Weekly Order during Mature Demand ②×MV	$50k	$80k	$70k	$200k
④	Amount on Order = Weekly Order × Lead Time: ③×①	$300k	$800k	$980k	$2080k
⑤	Percent Value of Part on Order: ④÷$2080k	14.4%	38.5%	47.1%	100%
⑥	On-order Material in Weeks of MV ④÷MV	1.5	4.0	4.9	10.4

Total Inventory Metrics during Mature Demand

		Weeks of Mature Demand	Dollars
①	RPI	7.6	$1520k
②	WIP	2.0	$400k
③	FGI	2.0	$400k
④	On-Hand Inventory: ①+②+③	11.6	$2320k
⑤	On-Order Material	10.4	$2080k
⑥	Committed Inventory: ④+⑤	22.0	$4400k
⑦	Consignment Inventory	1.5	$300k
⑧	Total Committed Inventory: ⑥+⑦	23.5	$4700k

IV-3 End-of-Life Considerations for Experiment 0

Total PCFT = $7,800,000. Net profit = $78,000(i/100), where i is the profit as a percent of PCFT.

The following table summarizes the impact on the profitability of various margins i.

Write-Off as a Function of Profit on Shipped Units

①	Profit Margin i	5%	10%	20%	30%
②	Profit from Trade Units $7.8M × ①	$390k	$780k	$1560k	$2340k
③	Leftover Material	$64,615			
④	Leftover Material as % of Net Profit: ③÷②	16.57%	8.28%	4.14%	2.76%
⑤	Consignment	$300,000			
⑥	Consignment as % of Net Profit ⑤÷②	76.92%	38.46%	19.23%	12.82%
⑦	Total EOL Material as % of Net Profit: (③+⑤)÷②	93.49%	46.75%	23.37%	15.58%

The following table shows the impact on Class C EOL material of reducing safety stock levels. These results were computed using means other than simulation.

Weeks of Class C Safety Stock	Class C EOL Material
16 weeks	$64,615
15 weeks	$35,385
14 weeks	$13,846
13 weeks	$0

IV-4 Why There Is Class C material Left Over for Experiment 0

The last period in which we expect to receive orders is week 68. The end of week 55 is 13 weeks before the end of the product life cycle. From the Adder order forecast in Fig. 2 on page 83 and the target RPI safety stock for class C material being 16 weeks of the 13-week leading average forecast (Table Ib on page 83), at the end of week 55 the amount of class C material in RPI should theoretically be 16/13 of the total demand to the end of life, or $(16/13) \times (1\frac{3}{4} \times V) = (28/13) \times V$ units, where V = 80.

In week 56, we need to start building the units for orders received in week 55. Ignoring the current FGI, the maximum new build from week 56 to the the end of life is equal to the demand from week 55 through the end of life, that is, 2V. Thus, at the end of week 55, there is more class C material on hand—enough to build (28/13) × V units—than needed for the demand to the the end of the product life cycle.

Remember that we did not consider units in FGI. If we want to reduce FGI units down to 0 by the end of the product life cycle, the total new build must be less than that computed above, and hence there will be even more class C material left over.

In summary, one reason for the leftover class C material is that the safety stock computation requires holding more class C raw material in RPI 13 weeks before the end of life than can be consumed by orders received in the last 14 weeks of the product life cycle.

IV-5 Why Orders Cannot Be More than 14 Weeks Late for Experiment 1a

Assume that an order comes in during week x. In the worst case we have not yet ordered any material for the unit that goes with this order. The earliest the material can be ordered is week x+1, and the longest lead time part will be delivered during week (x+1)+14, which is week x+15. Since build time is 2 weeks, the unit is ready in week x+17. Since transit time is 1 week, the unit is delivered to the customer in week x+18. Since the quoted availability is 4 weeks, on-time delivery means the customer should receive it in week x+4. This means that the lateness is 14 weeks.

Determination of near optimal stock levels for multi-echelon distribution inv...
Masters, James M
Journal of Business Logistics; 1993; 14, 2; ABI/INFORM Global
pg. 165

DETERMINATION OF NEAR OPTIMAL STOCK LEVELS FOR MULTI-ECHELON DISTRIBUTION INVENTORIES

by

James M. Masters
Ohio State University

Managers are under increasing pressure to reduce inventories at all levels. Lower inventory investment can lead to higher asset productivity and to greater return on investment, and so managers are searching for ways to reduce or eliminate stock. In a recent study of *Fortune 500* firms in the chemical, electronic, food, and pharmaceutical industries, Loar has reported that average inventory holdings across these four groups declined from 23% in the food industry to 37% in the chemical industry over the period from 1970 to 1987.[1] The apparent success of the "Just in Time" approach to inventory control in the production environment has sent management looking for ways to make similar dramatic reductions in post-production distribution inventories. Firms have developed "Quick Response" and "Direct Store Delivery" systems[2] to send frequent, small replenishment shipments to retail locations in an attempt to emulate the Just-in-Time approach in a retail setting. Some very impressive results have been reported for Quick Response systems. For example, a clothing manufacturer reported sales growth of 64% to 129% over a three-year period as a result of implementing a quick response inventory approach that greatly increased item availability at retail stockage points.[3] However, the inventory theoretic basis for these systems has not been well developed. How much total inventory is required in such a system? How should this quantity be allocated across locations? What is the appropriate trade-off between reduction in inventory and the increase in transportation costs which may result? How should these decisions be made?

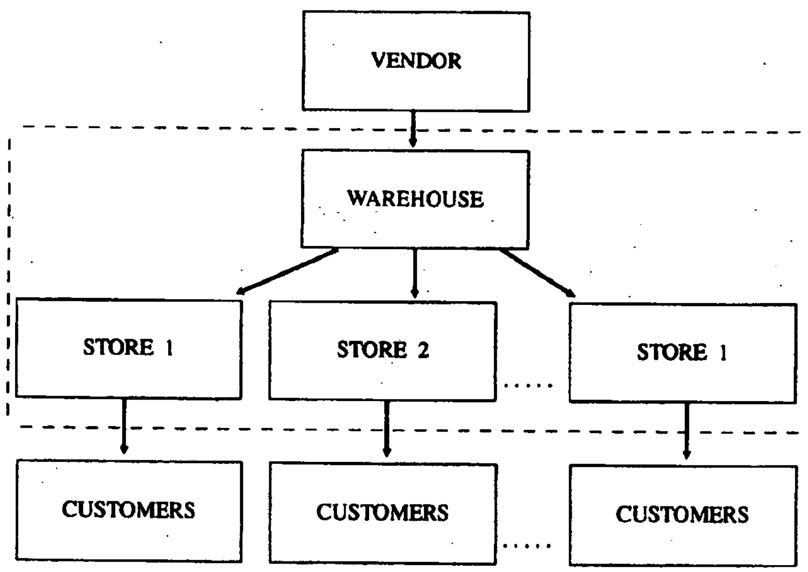
Several factors increase the complexity of this apparently straightforward problem. Many retail inventory systems involve hundreds, if not thousands, of individual items; an inventory decision must be reached for each item. In addition, most retail systems are multi-echelon in nature; that is, inventory is maintained at different levels within the system. For example, a retail store operation might involve a central warehouse that would order and receive property in bulk shipments

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from its suppliers. The warehouse would send smaller quantities to each individual store to support sales. Stockage decisions must be made both at the store and at the warehouse level. Stockouts at the stores would typically lead to lost sales, while stockouts at the warehouse would lead to backorders to the stores. These backorders, in turn, might lead to stockouts and lost sales at the stores. A diagram of such a system is included as Figure 1. The optimal allocation of inventory investment across items, locations, and echelons is far from obvious in such a situation. Another important consideration is the effect of the inventory system on transportation costs. Movement towards a system involving many, small, frequent replenishment shipments might increase transportation costs dramatically. In a common carriage situation, small shipments would typically move at much higher rates than the large shipments they would replace. In a private carriage operation, frequent replenishment might lead to a much larger number of vehicle movements, and hence to greater total transportation costs.

FIGURE 1

A MULTI-ECHELON INVENTORY SYSTEM



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This paper details the development of an analytic stockage model that determines the near optimal amount and distribution of stockage in a multi-item, multi-location, multi-echelon retail inventory/distribution system. The distribution environment includes items with high value and relatively low levels of stochastic demand, lost sales at the store level, backorders at the warehouse level, and one-for-one resupply logic. The analysis explicitly considers trade-offs between inventory holding costs, lost sale costs, and transportation costs.

Many different kinds of products would be good candidates for this form of inventory control. In men's clothing such as dress shirts and trousers, style, fabric, color, and size combinations result in a very large number of individual stock keeping units. A specific item—for example, a dress shirt in solid blue 100% cotton oxford cloth with button down collar, 15 inch neck, and 33 inch sleeve—might have total sales of no more than one or two per day at a given store. Similarly, auto parts stores must maintain a wide array of parts because of the proliferation of automobile make/model/year combinations that must be served. Even in a relatively high usage component such as car batteries, a typical store will sell no more than a handful of a specific SKU on any given day. Another example might be compact discs sold in music stores; at any given time only a handful of current releases are selling in large volume, while thousands of other SKUs sell at a rate of two or three per month per store. Many other such situations exist where retailers must maintain a broad range of individual items with low daily sales rates.

The remainder of the paper is organized as follows. Section 1 is a brief review of the inventory control literature which forms the theoretic basis for the model developed here. Section 2 is a more detailed discussion of Palm's theorem, which is the queueing theorem that lies at the heart of the model. Readers who are familiar with Palm's theorem can omit this section without loss of continuity. Section 3 contains the development of the model, while Section 4 describes the procedures that have been developed to solve the model; that is, to develop specific numerical solutions. Section 5 includes a discussion of validation issues. Section 6 includes a discussion of implementation issues as well as suggested areas for further development and study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LITERATURE REVIEW

Much inventory modeling in the distribution environment is in the tradition of economic lot sizing and statistical reorder point theory as has been described by Hadley and Whitin,[4] who recognized the pervasive presence of multi-echelon inventory systems throughout the business world and yet recommended the adoption of single location, single echelon models because of the complexity and intractability of the multi-echelon problem. Later work by Schwarz[5] and many others[6] has produced a large set of models that generally seek to identify optimal lot sizes and safety stocks in a multi-echelon framework. In addition, simulation models have been developed to describe the performance of multi-echelon inventory systems.[7] These simulations can capture the complex interactions of the multi-echelon problem, and can compare the performance of various inventory scenarios, but do not generate optimal inventory policies. These models are not well suited to the Quick Response retail environment because they tend to focus on large lot sizes with infrequent replenishment and they usually treat demand during stockout as backorders. In a Quick Response environment, no fixed lot size is imposed and stock is simply replenished in reaction to sales.

For over ten years, interest has grown in the practice of Distribution Resource Planning (DRP), as described by Whybark[8] and developed by Martin.[9] This technique applies the approach of Material Requirements Planning[10] to the problem of distribution inventories; that is, the inventory problem is addressed as a deterministic scheduling problem. In this basic framework, large order quantities are established, and no attempt is made to optimize formally the behavior of the inventory system. Attention is directed to the planning and scheduling of inventory network flow, while lot sizing and safety stock level determinations are considered secondary issues. Just as lot-sizing heuristics and the effects of uncertainty have been incorporated in MRP,[11] more recent studies[12] have involved the incorporation of lot-sizing logic in DRP. In practice, DRP generally involves high volume items and large lot sizes. In the same sense that Just-In-Time inventory systems can be thought of as MRP systems where the planning interval, or "time bucket," has been compressed from weeks to hours, Quick Response can be thought of as a kind of DRP system where the planning interval is on the order of a single day and no minimum shipment size is imposed. Unfortunately, DRP is not well suited to deal with the inherently stochastic nature of this situation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The model developed in Section 4 attacks a retail inventory problem by adapting and applying modeling techniques that have evolved over a considerable period of time but which have not yet been applied to retail problems. Feeney and Sherbrooke[13] established a basic model of item stockage with one-for-one replenishment based on a queueing theorem attributed to C. Palm[14] and developed performance measures for the backorder and lost sales cases that are outlined in Section 4. This modeling approach has been extensively developed and applied in the context of spare parts stockage for military aircraft. In this context, only the backorder situation was relevant. Subsequent development of the basic model was triggered by Sherbrooke,[15] who developed METRIC (Multi Echelon Technique for Recoverable Item Control). This model established the procedure of linking upper-echelon warehouse stockage to lower-echelon inventory performance through the mechanism of expected warehouse delay time. Further development included explicit modeling of sub-component relationships by Muckstadt[16] and extension to the case of time-dynamic solutions by Hillestad.[17] Theoretic work in this (s-1,s) paradigm continues to the present.[18]

PALM'S THEOREM AND INVENTORY CONTROL

An analytic approach to modeling inventory systems with one-for-one replacement can be based on a theorem published in 1943 by C. Palm. This theorem was first used in a queueing system context to describe the expected number of telephone calls that would be going on at the same time, and hence, to establish the number of trunk lines that would be needed. Since that time the theorem was seen to have a more general utility, particularly when applied to inventory situations.

Simply put, the theorem is as follows: in a queueing system where:

1. Arrivals occur as a Poisson process with a mean of λ, and
2. Service times are independent of the arrival process and are described by any stationary stochastic process with a mean of τ, and
3. A sufficient number of servers exist so that all servers are never busy simultaneously; the number in the system at any point in time is a Poisson distributed random variable with a mean of $\lambda\tau$.

The proof of this theorem follows from the notion that if a Poisson process is randomly censored, then the censored process is also Poisson.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As an illustration of the basic idea, consider a simple example of pedestrians crossing a bridge. Suppose that people arrive at a bridge randomly at a rate of 2 per minute (λ = 2/minute), and that this arrival process is accurately described by a Poisson distribution. Each walker immediately begins to cross the bridge; that is, the bridge is sufficiently wide so that no one must wait to begin crossing. Each person proceeds at his or her own chosen speed, and the average speed of all walkers is such that the average crossing time is 5 minutes (τ = 5 minutes). It follows from Palm's theorem that the average number of people on the bridge at any given time is 10 (2/minutes x 5 minutes), and that the probability that exactly N people are on the bridge at any given time is the probability that X=N in a Poisson distribution with a mean of 10.

This theorem can be directly applied to an inventory system with one-for-one replenishment and backordering of demand during stockouts by considering the demand process as the arrival process and treating the stock replenishment process as the service process. For example, if demand occurred according to a Poisson process at a rate of 2 units per day, and each demand resulted in an immediate reorder with an average replenishment time of 5 days, then by Palm's theorem the number of units in replenishment is Poisson distributed with a mean of 10; that is, Palm's theorem provides the exact distribution of the random variable, the quantity in replenishment. Given information about the initial stockage of the system, it is now possible to describe the performance of the inventory system in terms of item availability and expected backorders. To continue the current example, suppose 15 units of stock were initially provided. Given one-for-one replenishment, it follows that the total number of units on hand and on order will always equal 15. The quantity on hand, however, will be a random variable with values from 0 to 15. The probability that exactly 5 units, say, will be in stock at any given time is the probability that exactly 10 units are in replenishment. Thus Palm's theorem can be used to describe the distribution of on hand stock. Similarly, 4 units will be on backorder if and only if exactly 19 units happen to be in replenishment; thus Palm's theorem can be used to describe the backorder distribution as well. Similar procedures can be used to describe the performance of inventory systems where demand is lost rather than backordered during stockouts.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

MODEL DEVELOPMENT

The basic approach of the model presented in this section is to formulate and solve the inventory allocation problem as a nonlinear mathematical program. Customer demands and system resupply times are treated as random variables, and inventory system performance is calculated on an expected value basis.

Measures of Inventory System Performance

Assume that the customer demand per unit time for an item at a store is a discrete random variable from a stationary Poisson process. Then let:

λ = the demand rate of the item;

τ = the average resupply time to the store; the mean of an arbitrary, stationary, independent stochastic process;

q = the authorized inventory quantity; and

n = a random variable; the quantity in resupply; that is, the total replenishment quantity on order by the store.

Given that demand is independent of the resupply time and that inventory is reordered on a one-for-one basis as it is used, it has been shown by Feeney and Sherbrooke[19] that, by applying Palm's Theorem, the probability that a given quantity of n units is in resupply in the backorder case is:

$$P[n \mid \lambda,\tau] = \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \; ; \; n=0,1,\ldots\infty \tag{1}$$

and is in the lost sales case:

$$P[n \mid \lambda,\tau] = \frac{\dfrac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}}{\displaystyle\sum_{k=0}^{q} \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}} \; ; \; 0 \le n \le q \tag{2}$$

and:

$$P[n \mid \lambda,\tau] = 0 \; ; \; n>q \tag{3}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Note that the quantity in replenishment or resupply is based on a random variable from a Poisson distribution with a mean of $\lambda\tau$, and that this result is generally true without regard to the specific probability distribution which describes the resupply times. The exact probability distribution described in equations (1), (2), and (3) can be used to develop measures of the steady state behavior or performance of these inventory systems.[20] In a system where customer demands which occur during periods of inventory stockout are backordered, the expected backorder quantity would be:

$$E[B] = \sum_{n=q}^{\infty} (n-q) \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \tag{4}$$

This quantity represents the expected number of backorders outstanding at any random point in time. In a system where an item experiences lost sales rather than backorders, the fraction of item demand per unit time that is not lost, or the item fill rate, would be:

$$F = \frac{\sum_{k=0}^{q-1} \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}}{\sum_{k=0}^{q} \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}} \; ; \; q>0 \tag{5}$$

Therefore the expected lost sales per unit time could be expressed as:

$$E[L] = \lambda(1.0-F) \tag{6}$$

Feeney and Sherbrooke were also able to show that these results can be generalized to the case where the demand distribution is any compound Poisson distribution. This provides a much richer variety of distributions that can be used to model actual customer demand.

Since the use of a Poisson or compound Poisson distribution to represent customer demand is crucial to the development of the model, this assumption should be examined in any practical application to insure that the assumption is reasonable,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that is, that empirically collected demand data would be well described by such a distribution. In general, the Poisson is a reasonable choice for a demand distribution in the case where a relatively large number of customers all have an equal but small probability of purchasing an item at a given store during a given time interval and all customers act independently. In such a case the number of demands per time interval is a random variable with a binomial distribution. If the number of customers is large, and the probability associated with each customer is small, then the Poisson is an excellent approximation for this binomial distribution. The Poisson is also an excellent approximation even if all of the customer's purchase probabilities are not exactly equal, that is, even if some customers are more likely to purchase than others.[21] Poisson distributions are characterized by variance to mean ratios of 1.0; that is, the variance is equal to the mean. In practice, some demand distributions display more variance than this. The family of compound Poisson distributions, which includes, for example, the Negative Binomial distribution, can be used to describe the demand distribution in these cases. A compound Poisson process would naturally arise, for example, if customers arrive as a simple Poisson process and the number of units each customer purchases is a random variable. Thus a Poisson or compound Poisson distribution is a reasonable model for customer demand in many situations, but in any specific application empirical demand data should be used to test the assumption.

A Multi-echelon Model

This single item, single location model can be expanded into a multi-echelon formulation that will conform to a retail stockage problem as follows. Suppose that an item is stocked at a single warehouse to resupply a number of individual store inventories that support sales of the item at the retail level. Demands that may occur at a store while that location is stocked out become lost sales without exception; no demand is transferred between stores. No property is transshipped between stores, and no expediting of replenishment orders is permitted. Stock replenishment orders placed by the stores on the warehouse are backordered during intervals when the warehouse is stocked out. The warehouse itself reorders from a vendor and is resupplied, without backorder, after a random resupply time has elapsed. Further, suppose that the item is reordered at all locations, including the warehouse, on a one-for-one basis as it is used.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Let:

I	=	the total number of stores;
λ_i	=	the demand rate at store i;
τ_i	=	the mean resupply time from the warehouse to store i;
q_i	=	the stock quantity authorized at store i;
F_i	=	the item fill rate at store i;
λ_0	=	the total demand rate at the warehouse;
τ_0	=	the mean resupply time to the warehouse; and
q_0	=	the stock quantity authorized at the warehouse.

It follows that the total replenishment demand rate placed on the warehouse, λ_0, will be:

$$\lambda_0 = \sum_{i=1}^{I} F_i \lambda_i \tag{7}$$

Then the expected backorders at the warehouse can be estimated per equation (4) as

$$E[B_0] = \sum_{n_0=q_0}^{\infty} (n_0-q_0)P[n_0 \mid \lambda_0,\tau_0] \tag{8}$$

Backorders at the warehouse do not in themselves necessarily reduce service levels at the stores; as long as some stock remains at each store, no sales will be lost. However, as backorders accumulate at the warehouse it becomes more likely that stockouts will occur at the stores. In order to model this relationship at the store level, we can add an expected waiting time to the store's normal resupply time. Logically, this waiting time will be related to the number of outstanding backorders. This is equivalent to saying that when a store places a replenishment demand upon the warehouse while backorders exist, this new demand will not be satisfied until all prior backorders have been satisfied, that is, until this new order has worked its way to the top of the waiting list. A general rule of queueing theory known as Little's Law, "$L = \lambda W$," states that the average number of entities in a system will equal the arrival rate times the waiting time. By recognizing that the expected

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

backorder quantity, $E[B_0]$, represents the average number of entities in a queueing system and that λ_0 represents the arrival rate, we can use this rule to infer the system waiting time, or the average delay time due to backorders. That is, this backorder quantity implies that each store will observe an expected delay, or average waiting time due to backorders, in its resupply time which can be calculated as:

$$W_0 = \frac{E[B_0]}{\lambda_0} \qquad (9)$$

Adding this expected delay to the expected resupply time at each store leads to a modified expression for the mean of the random variable, the quantity in resupply to store i:

$$\mu_i = \lambda_i(\tau_i+W_0) \qquad (10)$$

Given this mean value, the item fill rate at store i can be estimated as in equation (5) as:

$$F_i = \frac{\sum_{k=0}^{q_i-1} \frac{e^{-\mu_i}\mu_i^k}{k!}}{\sum_{k=0}^{q_i} \frac{e^{-\mu_i}\mu_i^k}{k!}} \; ; \; q_i>0 \qquad (11)$$

And thus the total lost sales (per unit time) across all stores can be calculated as:

$$\sum_{i=1}^{I} E[L_i] = \sum_{i=1}^{I} \lambda_i(1.0-F_i) \qquad (12)$$

Incorporating the mean warehouse backorder delay time as a component of the mean resupply time to the store directly ties the performance of the warehouse to the performance of the stores, but it also leads to the violation of an assumption of the model, namely, that the demand processes and the resupply processes are stochastically independent. Thus the model formulation is only approximate, and model results must be tested to determine the accuracy of the approximation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Inclusion of Multiple Items and Inventory Costs

This multi-echelon model can be expanded to include investment trade-offs across a range of items by considering the holding costs and expected shortage costs associated with various stockage policies. The purpose of such a model is to determine the optimal range and depth of item stockage across items and locations.

Let:

J	=	The total number of items in the system;
C_j	=	Cost of item j, in dollars per unit;
V_j	=	Lost Sale Cost of item j, in dollars per unit;
$E[L_{ij}]$	=	Expected Lost Sales at location i of item j, per unit time;
H	=	Holding Cost Fraction, in dollars per dollar held per unit time;
q_{ij}	=	Stock Allocation of item j at location i.

Given that there is no demand substitution between items when stockouts occur, then $E[L_{ij}]$ can be computed independently for each item in the system. With known lost sale costs for each item, the inventory allocation decision can be formulated as an unconstrained nonlinear optimization problem with integer valued decision variables. One would seek to minimize the sum of the expected lost sale costs and inventory holding costs per time period;

$$\text{Minimize: } \sum_{j=1}^{J}\left[\sum_{i=1}^{I} E[L_{ij}]V_j + HC_j q_{ij}\right] + HC_j q_{0j} \tag{13}$$

For simplicity, in this formulation holding costs are charged against all items on hand or on order. Given continuous review and one-for-one reordering, the total amount of stock on hand plus on order is a constant for any given stockage allocation. In the (frequently encountered) case where true lost sale costs are unknown, the inventory allocation decision can be formulated as a nonlinear optimization problem with a linear budget constraint. One would seek to minimize the sum of the lost sales across all items, as modified by any appropriate importance weighting W_j, subject to a budget limit on total inventory investment, say B dollars:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\text{Minimize: } \sum_{j=1}^{J} \sum_{i=1}^{I} W_j E[L_{ij}] \tag{14}$$

$$\text{Subject to: } \sum_{j=1}^{J} \sum_{i=0}^{I} C_j q_{ij} \leq B \tag{15}$$

In each case (that is, known or unknown lost sale costs) the problem formulation is separable with respect to the decision variables (the q_{ij}), and the objective function is nearly convex. As a result, optimal solutions to the model formulations can be derived using straightforward marginal analysis techniques[22] as described in Section 5.

Incorporation of Lot Sizing at the Warehouse

The model as described thus far is based on the policy that inventory is reordered on a one-for-one basis as it is sold at or sent to the stores. Thus the lot size is essentially predetermined, and the inventory stockage decision actually involves the amount and distribution of safety stocks in the system. This policy of one-for-one replenishment is common and is optimal at the store level for high value, low volume items.[23] However, it may be less realistic to assume that the warehouse, which assembles demand from all stores, would itself reorder one-for-one. A simple heuristic extension can be made to the model which allows the incorporation of a warehouse lot size as a model parameter. A more complete treatment might allow this lot size as a decision variable; in practice, however, lot sizes are often dictated by externalities such as vendor minimum order size policies. Given a predetermined lot size, say l_0, and a total stockage authorization of q_0, the warehouse would place a replenishment order for l_0 units when the reorder point, r_0, is reached, where:

$$r_0 = q_0 - l_0 \tag{16}$$

In this situation the expected backorder quantity would be calculated as:

$$E[B_0] = \sum_{n_0=r_0}^{\infty} (n_0 - r_0) P[n_0 \mid \lambda_0, \tau_0] \tag{17}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Unlike the case of one-for-one replenishment, this quantity is not the average number of outstanding backorders at a random point in time; rather, this is the expected number of outstanding backorders at the end of an order cycle. If the lot size is sufficiently large and the vendor resupply time is sufficiently brief, so that in general at most one replenishment order exists at any point in time, then we can approximate the average number of outstanding backorders at any point in time as the average number of backorders during the stockout interval, $E[B_0]/2$, times the proportion of the cycle time that is the stockout interval, which can be expressed as $E[B_0]/q_0$ (see Figure 2). Thus the expected delay time due to stockouts at the warehouse can be approximated as:

$$W_0 \approx \frac{E[B_0]^2}{2q_0\lambda_0} \tag{18}$$

Using this value for W_0 in the basic model can therefore accommodate the situation where the warehouse reorders with a fixed lot size. The quality of this approximation is discussed in Section 5.

Inclusion of Transportation Cost Trade-offs

A more complete treatment of the decision would recognize that many transportation alternatives exist, so that faster or slower replenishment times could be selected, and that generally speaking, faster replenishment times are more expensive than slower replenishment times. All other things equal, reducing replenishment times will reduce the desired or optimal inventory stockage in the system, but at a cost. The correct decision is to choose the replenishment speed and inventory level jointly so that the sum of the inventory holding, lost sale, and transportation costs are minimized.

In terms of the model, this means treating τ_i as a decision variable rather than as a parameter of the problem. In general, each τ_i could be individually varied by a manager by choice of carrier, air freight, overnight service, and so forth. Treating each store's mean replenishment time as a decision variable is conceptually possible, but it leads to a very unwieldy problem formulation. A more straightforward approach would be to determine a system-wide decision variable, $\bar{\tau}$, which would

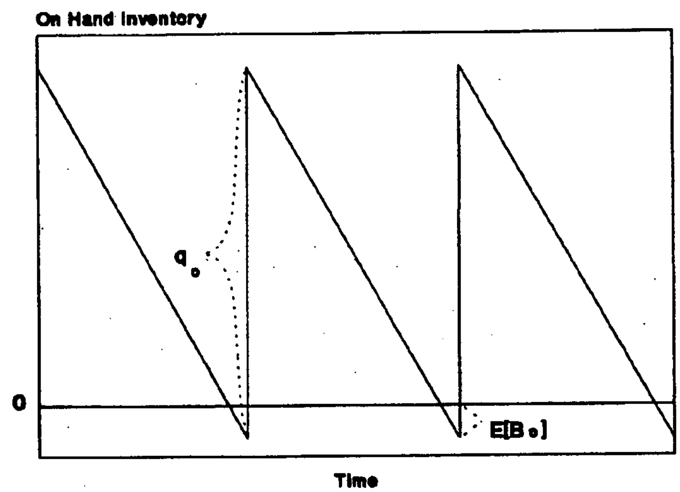
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

be the average replenishment time across all stores. Reducing the system replenishment time would result in an increase in T, the total transportation cost per time period. In general, the transportation cost $T[\tau]$, is a monotonically decreasing function of the average replenishment time. Given that a choice of $\bar{\tau}$ leads to an unambiguous definition of each, τ_i, inclusion of the transportation decision in the problem framework can be accomplished by modifying equation (13):

$$\text{Minimize: } \sum_{j=1}^{J} \left[\sum_{i=1}^{I} E[L_{ij}]V_j + HC_j q_{ij}\right] + HC_j q_{0j} + T[\bar{\tau}] \qquad (19)$$

FIGURE 2

THE WAREHOUSE REPLENISHMENT CYCLE



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

SOLUTION PROCEDURES

Numerical solution of this multi-item, multi-echelon problem will involve four steps:

1. The determination of the system performance as defined in equation (12) for any given stockage allocation; that is, the system fill rate which will result from a given set of q_{ij} for a given item.
2. The determination of the optimal distribution of a given quantity of stock of an item across the stockage locations; that is, for a given total quantity of an item, say Q_j, the identification of the values of q_{ij} which will:

$$\text{Minimize: } \sum_{i=1}^{I} E[L_{ij}] \tag{20}$$

$$\text{Subject to: } \sum_{i=0}^{I} q_{ij} = Q_j \tag{21}$$

3. The allocation of stockage across items so as to minimize total costs, as in equation (13), or so as to minimize the weighted sum of lost sales subject to an investment constraint as in equations (14) and (15).
4. Simultaneously determining the appropriate stockage allocation and mean resupply time so as to minimize the sum of inventory holding costs, lost sale costs, and transportation costs as in equation (19).

Procedures to accomplish these four steps are as follows:

Determination of System Performance

In a typical "METRIC like" multi-echelon model,[24] shortages are backordered at all levels in the system. As a consequence, performance can be estimated with a simple recursion. Given an allocation of stock to the upper level, expected backorders at the upper level are calculated, and delay time is computed. This delay time is added to the resupply time at each lower level stocking point, and the system performance at the lower level is estimated. This procedure is appropriate

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in a pure backorder system because the total demand placed on the upper level location is simply the sum of the demands observed at the lower level locations. Total demand is conserved, and no demands are ever lost. In the model described above, however, this property is lost. Demand at the warehouse, and hence the warehouse delay time, are functions of the fill rates at the stores, and the store fill rates are themselves functions of the warehouse delay time. We, therefore, use an iterative scheme to converge to a simultaneous solution for the fill rates and delay time implied by any allocation of q_{ij} as follows:

1. Set an initial estimate of the warehouse demand rate to be the sum of the demand rates across the stores:
2. Estimate warehouse backorders and delay time using this estimate of the warehouse demand rate in equations (8) and (9).
3. Estimate the store fill rates using these results in equations (10) and (11).

$$\lambda_0' = \sum_{i=1}^{I} \lambda_i \tag{22}$$

4. Recompute the warehouse demand rate using the fill rates from step 3 in equation (7).
5. Repeat steps 2 through 4 until λ_0 converges.

Computational experience has shown that this procedure typically converges within five to twenty iterations. However, in some situations this procedure will cycle indefinitely between two sets of estimates, each set returning the other. An alternative procedure that is somewhat slower, but has never failed to produce convergence in our experience, is to replace the old estimate of λ_0 not with the new estimate, but rather with the average of the old and the new estimate. A reasonable compromise between algorithm speed and certainty of convergence is achieved by using the first procedure (replacing the old estimate with the new estimate) for up to forty iterations, and then switching to the second procedure (averaging the old and new estimates).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Determination of the Optimal Distribution of Stock for an Item

In most realistically sized problems there will be a very large number of unique allocations of a given amount of stock of an item across the locations that might be considered, as is demonstrated in Table 1. Enumeration and evaluation of each of these allocations would not be efficient or desirable.

A simple approach to this problem is to perform an incremental analysis of the allocation decision. This seems plausible since the performance functions defined in (1) and (2) are convex with respect to an increase in stockage. Given that Q_j assets have been allocated across I locations, an allocation of Q_j+1 assets can be found by considering and evaluating the I ways to augment the existing allocation by placing one additional asset at a location. The allocation that produces the best performance is selected as the correct allocation of Q_j+1 assets. By starting with $Q_j = 0$, this procedure can allocate any number of assets, and it will generate all the intermediate solutions; that is, the suggested allocations for $0,1,2,...Q_j$ assets. This is a useful feature since the correct value for Q_j is not generally known in advance. In general, Q_jI performance evaluations must be made to determine the allocation of Q_j assets at I locations using this incremental approach.

TABLE 1

THE NUMBER OF POSSIBLE ALLOCATIONS OF Q_j ASSETS ACROSS 10 AND 20 LOCATIONS

Total Number of Assets [Q_j]	Allocations Across 10 Locations	Allocations Across 20 Locations
10	$9.2378 \cdot 10^4$	$2.0030 \cdot 10^7$
20	$1.0015 \cdot 10^7$	$6.8923 \cdot 10^{10}$
30	$2.1192 \cdot 10^8$	$1.8852 \cdot 10^{13}$
40	$2.0545 \cdot 10^9$	$1.3973 \cdot 10^{17}$
50	$1.2566 \cdot 10^{10}$	$4.6253 \cdot 10^{17}$
100	$4.2634 \cdot 10^{12}$	$4.9104 \cdot 10^{23}$

The simple approach described above is very fast, but it will not always generate the optimal allocation. Due to the interaction between echelons, situations exist

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

where the optimal allocation of Q_j+1 assets does not consist of the optimal allocation of Q_j assets with one more asset added at one location; rather, adding one asset to Q_j permits a new optimal allocation which is considerably different from the prior optimal allocation. In practice, this situation will occur when stores have identical or nearly identical values of $\lambda_i \tau_i$. A slightly more complicated marginal analysis that first partitions assets between the echelons will overcome this difficulty. There are Q_j+1 ways to partition Q_j assets between the warehouse and the stores (0 at the warehouse, Q_j at the stores; 1 at the warehouse, Q_j-1 at the stores; etc.). For each such partition, use the simple incremental approach to develop the allocation across the stores. Comparison of performance evaluations across the partitions will produce the optimal allocation. This technique will require more performance evaluations; in general:

$$\text{Number of Evaluations} = I \left[\frac{Q_j(Q_j+1)}{2}\right]+1 \tag{23}$$

As is shown in Table 2, this echelon partitioning approach can considerably increase the number of performance evaluations required.

TABLE 2

COMPARISON OF SEARCH HEURISTICS

Number of Assets	10	10	100	100
Number of Locations	10	20	10	20
Number of Possible Allocations	$9.2 \cdot 10^4$	$2.0 \cdot 10^7$	$4.2 \cdot 10^{12}$	$4.9 \cdot 10^{23}$
Evaluations with Incremental Allocation	100	200	1,000	2,000
Evaluations with Echelon Partitioning	560	1,120	50,510	101,020

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Optimal Inventory Budgets and Budget Allocations

Given that we can determine the optimal distribution of any quantity of a given item across the stockage locations, the next step is to determine the optimal level of inventory investment for the system as a whole, that is, the total quantity and deployment plan for each of the J items in the system. In the case where holding costs and lost sale costs are known, this consists of solving the unconstrained minimization problem given in (13). In the case where lost sale costs are not known, this consists of solving the constrained minimization problem given in equations (14) and (15). The constrained problem can be replaced with a generalized unconstrained Lagrange function:[25]

$$\text{Minimize: } \sum_{j=1}^{J} \sum_{i=1}^{I} W_j E[L_{ij}] - \theta \sum_{j=1}^{J} \sum_{i=0}^{I} C_j q_{ij} \tag{24}$$

which is structurally equivalent to equation (13). Solutions to equation (24) are efficient and undominated solutions to the problem presented in (14) and (15). For every investment budget, B, there is a multiplier, θ such that the solution to equation (24) is a stockage allocation such that no other allocation could provide a higher level of performance at an equal or lower total investment in inventory.

Solving equation (24) requires finding the value of θ, which minimizes the weighted sum of lost sales without violating the budget constraint. Note that θ can be interpreted as the implied benefit to cost ratio of the incremental unit of each item, where the benefit is the reduction in weighted lost sales provided by the last unit of stock. At an optimal solution, the marginal reduction in weighted lost sales per dollar of unit cost will be approximately equal across all J items in the system. A simple solution procedure is therefore as follows:

1. For each item, for each total stock quantity Q_j, calculate a marginal benefit to cost ratio, R_{Qj}, as follows:

$$R_{Qj} = \frac{\left(\sum_{i=1}^{I} W_j E[L_{ij} | Q_j]\right) - \left(\sum_{i=1}^{I} W_j E[L_{ij} | Q_j + 1]\right)}{C_j} \tag{25}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

using the performance optimizing allocations of Q_j as identified above. This value represents the improvement in the weighted sum of expected lost sales per dollar of investment which would occur if stockage of item j were increased from Q_j to Q_j+1. Continue to compute these ratios until the marginal improvement becomes arbitrarily close to zero.

2. In a multi-echelon problem, it will sometimes occur that:

$$R_{Qj+1} > R_{Qj} \tag{26}$$

That is, occasionally the Q_j+1 asset has a greater contribution to reduction in weighted lost sales than does the Q_j unit. In such cases, preserve the convexity of the list produced in step 1 by combining the contribution of the two assets, and calculate the benefit to cost ratio of the two asset increment:

$$R_{Qj} = \frac{(\sum_{i=1}^{I} W_j E[L_{ij}|Q_j]) - (\sum_{i=1}^{I} W_j [L_{ij}|Q_j+2])}{2C_j} \tag{27}$$

Continue to combine asset increments until the list is convex; that is, until:

$$R_{Qj} < R_{Q'j}; \text{ for all } Q_j > Q'_j \tag{28}$$

3. Merge the "lists" of R_{qj}'s formed for each item in step 2 into a single, system-wide list, and sort the list in descending sequence. This list represents the order in which each asset increment will enter an optimal solution.
4. Proceed to a solution by moving down the system list in sequence, adding each additional asset increment to the solution until the addition of the next increment would exceed the budget, that is, until equation (15) would be violated.

This so-called "shopping list" approach[26] reduces a complex, non-linear, integer programming problem to little more than a system sort. In the case where lost sales costs and holding costs are known the solution procedure is very nearly the same. The W_j values used to compute the R_{Qj} values in equation equation (25)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in step 1, would be the lost sale costs, or V_j values. Step 4 would consist of moving down the list so long as including the asset increment would continue to reduce the value of the objective function defined in equation (13), and stopping when the next asset increment would cause an increase in these total system costs.

Transportation Cost Trade-off Analysis

Given that we can determine the appropriate total inventory and inventory allocation across items and locations for a given set of τ_i, or mean replenishment times, the final step is to treat these replenishment times as variables and solve the problem formulate in equation (19), that is, minimize the sum of the inventory holding, lost sale, and transportation costs.

If only a few transportation alternatives are under consideration, this is not a very difficult problem. Suppose, for example, that the manager wished to choose between common carrier LTL, a ground package service such as UPS, and an air service such as Federal Express. Given the transportation costs and replenishment times implied by each transportation option, one would simply solve the original problem three times, each time with a different set of transportation parameters, and choose the low cost alternative.

A more comprehensive approach to a solution would recognize that there are in fact a great many possible choices for the system mean replenishment time and would seek to establish $\bar{\tau}^*$, or the optimal system-wide mean replenishment time. This situation is depicted in Figure 3. In this illustration the dotted line labeled "Transport Cost" represents the total transportation cost per unit time as determined by choice of mean replenishment speed, that is, it is a graph of $T[\bar{t}]$. The dashed line labeled "Inventory Costs" represents the minimized sum of inventory holding costs and lost sale costs given the specific value of $\bar{\tau}$. The solid line labeled "Total System Cost" is the sum of the transportation cost, inventory holding cost, and lost sale cost. Solving the general problem involves finding the value of $\bar{\tau}$ and the associated stockage allocation which minimize this total cost curve.

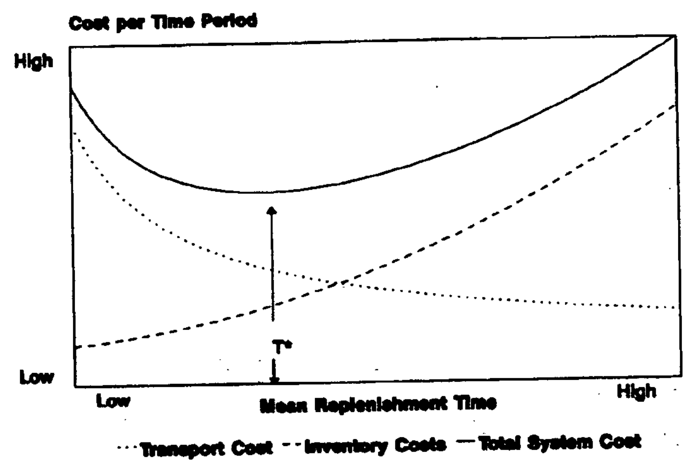
An efficient search procedure to find this solution is as follows. First, establish a meaningful increment in $\bar{\tau}$. That is, determine what constitutes a managerially meaningful difference in the mean replenishment time. Differences on the order of one day, or perhaps even a few hours, might be meaningful; differences on the order of one minute generally are not. This implies that while it is mathematically

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

possible to evaluate total systems costs at $\bar{\tau}$ = 3.0 days and at $\bar{\tau}$ = 3.01 days, in practice this comparison will not mean very much because the increment in $\bar{\tau}$ is so small. Once a meaningful increment is established, the search procedure will consist of sequentially evaluating the total cost expression as $\bar{\tau}$ is incremented from a low starting point. As $\bar{\tau}$ is increased, total system costs will gradually fall until the minimum is reached, and will then begin to rise.

FIGURE 3

INCORPORATING TRANSPORTATION SPEED AND COST IN THE INVENTORY DECISION



In spite of the selection of a meaningful increment in $\bar{\tau}$ it might seem that this search procedure will be a very lengthy process because it seems to imply that the entire basic problem must be resolved scores of times with new replenishment time parameters. In fact, however, this need not be the case. By starting at the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

low end of the range of feasible values for $\bar{\tau}$, this search procedure can take advantage of the nature of the stockage algorithm. Recall that the stockage algorithm starts with a given set of replenishment time parameters, establishes minimum stockage at all locations, and builds inventory incrementally to an optimal allocation. When we increment $\bar{\tau}$ by, say, Δ, it will generally be the case that the optimal stockage allocation given $\bar{\tau}+\Delta$ will consist of the optimal stockage allocation given $\bar{\tau}$ plus some additional stock at some locations. Thus each time we increment $\bar{\tau}$, we need not set the stockage allocation algorithm "back to zero." The analysis of $\bar{\tau}+\Delta$ can begin with the stock position identified in the analysis of $\bar{\tau}$. This approach to the transportation cost trade-off analysis will greatly reduce the computational effort required; nevertheless, on a large system with many items at many locations this may remain a lengthy job. On the other hand, this is not the kind of analysis that will be done on a daily basis. Establishing the optimal system replenishment time would probably be done no more often than quarterly.

MODEL VALIDATION

Estimation of System Performance

In order to test the accuracy of the inventory system model defined by equations (7) through (12), an event-oriented Monte Carlo simulation of the inventory system was constructed in FORTRAN. Sample means of system fill rate performance generated by the simulation were used as benchmarks to gauge the ability of the analytic model to estimate expected system fill rate performance associated with a given allocation of stock.

Performance comparisons on a sample item to be stocked at ten stores are presented in Tables 4 and 5. Demand rates and resupply times at each location are as shown in Table 3. System fill rates are averages calculated over 10 simulation runs, where each run was long enough to generate 20,000 demands. Given that the demand rate of 2.0 represented a rate of two demands per day, this would be the equivalent of simulating the operation of the inventory system over a period of 10,000 days.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 3

SAMPLE ITEM PARAMETERS

Location	Demand Rate	Resupply Time
Store i	2.0	1
Warehouse	—	4

One-for-one Reordering at All Locations

A set of results for the sample item is shown in Table 4, which compares the system fill rate calculated by the analytic model to the estimates generated by the simulation across a range of total stockage. In each case, the allocation of stockage across locations which was tested is a near-optimal distribution as determined by the analytic model.

Lot Sizing at the Warehouse

A second set of results for the sample item (in Table 5) examines the accuracy of the approximate treatment of lot-sizing at the warehouse provided by equation (18). An allocation of stock for the sample item was considered which included 70 units stocked at the warehouse and 6 units stocked at each of ten retail locations. Lot sizes of 1 through 25 units per replenishment order were examined. Representative results are displayed in Table 5; as would be expected, incorporating batched replenishment at the warehouse reduces the level of system fill rate which can be attained from a given allocation of stock. The heuristic treatment of batching provided by equation (18) produces model estimates which are reasonably accurate, but not so accurate as the situation with one-for-one replenishment at all locations.

IMPLEMENTATION ISSUES AND CONCLUDING COMMENTS

Like all modeling efforts, this analysis includes a number of limitations and simplifying assumptions which may or may not be reasonable in a given application. In general, the quality of a solution produced by the model will depend on the accuracy of all the model parameters, but the greatest potential problem is most

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE 4

ACCURACY OF THE INVENTORY MODEL
WITH ONE-FOR-ONE REPLENISHMENT

System Stock		System Fill Rate (%)		
Warehouse Stock	Store Stock	Model Estimate	Simulation Mean	Standard Error
50	40	75.213	75.216	.019
60	40	81.674	82.095	.129
60	50	87.400	87.324	.118
70	50	92.186	92.372	.059
70	60	95.764	95.447	.104
70	70	97.969	97.650	.059
80	70	99.254	99.098	.036

TABLE 5

ACCURACY OF THE INVENTORY MODEL WITH
LOT-SIZING AT THE WAREHOUSE

	System Fill Rate (%)		
Lot Size	Model Estimate	Simulation Mean	Standard Error
1	95.764	95.447	.054
2	92.443	95.366	.084
3	93.190	95.182	.072
4	93.567	94.999	.070
5	93.754	94.637	.104
10	93.598	93.412	.063
15	92.802	92.136	.132
20	91.775	90.492	.110
25	90.618	88.766	.113

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

likely to occur with the estimates of item demand rates (the λ_{ij}). First, there is the general problem of forecasting disaggregated item demand. In addition, the model assumes that item demand processes are mutually independent.

Demand Forecasting

Implementation of this technique would require a set of demand rate estimates, that is, a λ_{ij} estimate for each item at each store. Typically these data would be generated by a short range time series forecasting technique such as exponential smoothing. A complicating factor which will frequently arise, particularly in a consumer retail environment, is the fact that the firm's "demand" history is actually a record of sales. In many cases no records of lost sales are maintained; in the typical "self-service" retail environment, the inventory system has no way to recognize or record the occurrence of a lost sale. In such a system, use of the sales history data would generate a biased forecast of the true or latent demand for the item. Without capturing actual lost sales data (which may be very difficult to do), this can be a difficult problem to resolve, but it should not be ignored. A simple and approximate treatment would be to adjust the recorded sales data based on the achieved availability of the item. For example, if an item experienced sales of S_{ij} units at a given store during a given sales period, and if the item were in stock and available for sale for a fraction of time, A, of the sales period, then a rough estimate of the true demand during the period, D_{ij}, would be:

$$D_{ij} = \frac{S_{ij}}{A} \tag{29}$$

Correlated Demand

Another potentially serious implementation problem is the assumption that the item demand processes are independent across the stores. It is very easy to envision occasions where this would not be the case; for example, situations which might cause an unusually high demand for an item at one store (such as unusual weather) might cause high demand across all the stores. The model developed here depends strongly on the assumption that demands are uncorrelated. In the case where demands are correlated, more total system stock will be required to achieve a given fill rate, and a higher proportion of total system stock will generally be allocated to the warehouse than would be the case in the situation with uncorrelated demand. We are currently developing an extension to the approach presented here that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

explicitly considers demand patterns that are intercorrelated across time and across locations.

CONCLUSION

This paper has developed a model to determine near optimal inventory allocations in a multi-item, multi-echelon, distribution inventory system with backorders at the upper echelon and lost sales at the lower echelon. The model is appropriate where retail customer demands are independent and can be adequately represented by Poisson or compound-Poisson probability distributions, where resupply times follow any stationary, independent probability distribution, and where stock is replenished on a one-for-one basis as it is used. In addition, an approximate treatment was suggested to accommodate batch replenishment at the upper echelon, and the model was extended to include explicit consideration of the appropriate trade-off between inventory levels and transportation costs.

The distribution strategy described by this model is applicable across many different kinds of products, including appliances, electronics, auto parts, and clothing. The essential characteristics of the system are that a broad range of SKUs are maintained at a number of locations, that items exhibit low demand rates at any given location, and that stockouts result in lost sales.

This model and solution procedures were developed in an attempt to build a theoretic framework as well as a practical technique to model the "Quick Response" inventory deployment systems which many retailers are currently establishing. This approach captures the inherently stochastic nature of the problem in an optimum seeking framework. Additional work is underway to understand and incorporate the effects of intercorrelated demand structures which are frequently found in such systems.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

NOTES

[1]T. Loar, "Patterns of Inventory Management and Policy: A Study of Four Industries," *Journal of Business Logistics* 13, no. 2 (1992): 69-96.

[2]B. Knill, ed., "Quick Response: Now for the Hard Part," *Material Handling Engineering* 45 (March 1990): 67-78; and E. J. Muller, "Quick Response Picks Up Pace," Distribution 89, no. 6 (June 1990): 38-42.

[3]J. H. Hammond, "Coordination in Textile and Apparel Channels: A Case for 'Virtual' Integration," in *Proceedings of the Twentieth Annual Transportation and Logistics Educators Conference* (Columbus: Ohio State University Transportation and Logistics Research Fund, 1991), p. 124.

[4]G. Hadley and T. M. Whitin, *Analysis of Inventory Systems* (Englewood Cliffs, N.J.: Prentice-Hall, 1963).

[5]L. B. Schwarz, ed., *Multi-level Production/Inventory Control Systems: Theory and Practice* (Amsterdam: North-Holland, 1981).

[6]A. J. Clark, "An Informal Survey of Multi-echelon Inventory Theory," *Naval Research Logistics Quarterly* 19 (1972): 621-650.

[7]See, for example, T. D. Clark, R. E. Trempe, and H. E. Trichlin, "Complex Multi-echelon Inventory System Management Using a Dynamic Simulation Model," *Decision Sciences* 14, no. 3 (July 1983): 389-407.

[8]D. C. Whybark, "MRP: A Profitable Concept for Distribution," in *Proceedings of the Fifth Annual Transportation and Logistics Educators Conference* (Columbus: Ohio State University Transportation and Logistics Research Fund, 1975), pp. 82-93.

[9]A. J. Martin, *Distribution Resource Planning* (Englewood Cliffs, N.J.: Prentice-Hall, 1983).

[10]J. Orlicky, *Material Requirements Planning* (New York: McGraw Hill, 1975).

[11]B. J. Coleman and M. A. McKnew, "An Improved Heuristic for Multi-level Lot Sizing in Material Requirements Planning," *Decision Sciences* 22, no. 1 (Winter 1991): 136-156; and D. C. Whybark and J. G. Williams, "Material Requirements Planning Under Uncertainty," *Decision Sciences* 7, no. 4 (October 1976): 595-606.

[12]J. H. Bookbinder and D. B. Heath, "Replenishment Analysis in Distribution Requirements Planning," *Decision Sciences* 19, no. 3 (Summer 1988): 477-489;

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABOUT THE AUTHOR

James M. Masters is assistant professor of logistics management at Ohio State University. He holds a B.A. in English literature from the State University of New York at Buffalo, an M.S. in logistics management from the Air Force Institute of Technology, and an M.B.A. and Ph.D. in business administration from Ohio State. He is a member of the Society of Logistics Engineers, the Decision Sciences Institute, The Institute of Management Sciences, the Production and Operations Management Society, and the Warehousing Education and Research Council.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and R. L. Bregman, "Enhanced Distribution Requirements Planning," *Journal of Business Logistics* 11, no. 1 (1990): 49-68.

[13]G. J. Feeney and C. C. Sherbrooke, "The (s-1,s) Inventory Policy Under Compound Poisson Demand," *Management Science* 12, no. 5 (January 1966): 391-411.

[14]C. Palm, "Intensitatsschawankungen im Fernsprechverkehr," *Ericsson Technics*, no. 44 (1943): 1-189, as cited in G. B. Crawford, *Palm's Theorem for Nonstationary Processes*, Report R-2750-RC (Santa Monica, Calif.: Rand Corp., October 1981).

[15]C. C. Sherbrooke, "METRIC: A Multi-Echelon Technique for Recoverable Item Control," *Operations Research* 16, no. 1 (1968): 122-144.

[16]J. A. Muckstadt, "A Model for a Multi-Item, Multi-Echelon, Multi-Indenture Inventory System," *Management Science* 20, no. 4 (1973): 472-481.

[17]R. J. Hillestad, *Dyna-METRIC: Dynamic Multi-Echelon Technique for Recoverable Item Control*, R-2785-AF (Santa Monica, Calif.: Rand Corp., 1982).

[18]E. Smeitink, "Operating Characteristics of the (S-1,S) Inventory System with Partial Backorders and Constant Supply Times," *Management Science* 36, no. 11 (1990): 1413-1414.

[19]Same reference as Note 13.

[20]Same reference as Note 13.

[21]Same reference as Note 14 at p. 10.

[22]H. Everett, "Generalized Lagrange Multiplier Method for Solving Problems of Optimal Allocation of Resources," *Operations Research* 11, no. 3 (1965): 399-417; and B. L. Fox, "Discrete Optimization via Marginal Analysis," *Management Science* 13, no. 3 (1966): 210-216.

[23]Same reference as Note 4 at p. 204.

[24]Such as those described by sources at Notes 15, 16, and 17.

[25]Same reference as Note 22.

[26]T. J. O'Malley, *The Aircraft Availability Model: Conceptual Framework and Mathematics* (Bethesda, Md.: Logistics Management Institute, 1983).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Organization TC 3600 Bldg./Room KNOX

U. S. DEPARTMENT OF COMMERCE
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450
IF UNDELIVERABLE RETURN IN TEN DAYS

OFFICIAL BUSINESS

AN EQUAL OPPORT

RETURNED TO SENDER

UNDELIVERABLE AS ADDRESSED
UNABLE TO FORWARD

RECEIVED
JUL 27 2005
USPTO MAIL CENTER

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.